MAIL ROOM JAN 7 1997 90 PAT. & TRADEMARK OFF.

08/779460

PATENT

Attorney Docket No.: U 011098-6

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 20231

NEW APPLICATION TRANSMITTAL

Transmitted herewith for filing is the patent application of Inventors:

1-00 1. OSCAR JOHANNES MARIA **GODDIJN**
2. TEUNIS CORNELIS **VERWOERD**
3. RONNY WILHELMUS HERMANUS HENRIKA **KRUTWAGEN**
4. ELINE **VOOGD**

ETAL

WARNING: *Patent must be applied for in the name(s) of all of the actual inventor(s). 37 CFR 1.41(a) and 1.53(b).*

For (title):
ENHANCED ACCUMMULATION OF TREHALOSE IN PLANTS

1. Type of Application

This new application is for a(n) (check one applicable item below):

☑ Original (nonprovisional)

☐ Design

☐ Plant

WARNING: ***Do not*** *use this transmittal for a completion in the U.S. of an International Application under 35 U.S.C. 371(c)(4) unless the International Application is being filed as a divisional, continuation or continuation-in-part application.*

WARNING: ***Do not*** *use this transmittal for the filing of a provisional application.*

CERTIFICATION UNDER 37 CFR 1.10

I hereby certify that this New Application Transmittal and the documents referred to as enclosed therein are being deposited with the United States Postal Service on this date **January 7, 1997 in an envelope as "Express Mail Post Office to Addressee" Mailing Label Number EG130134065 addressed to the: Assistant Commissioner of Patents, Washington, D.C. 20231**

Geraldine Marti
(type or print name of person mailing paper)

Geraldine Marti

(Signature of person mailing paper)

NOTE: Each paper or fee referred to as enclosed herein has the number of the "Express Mail" mailing label placed thereon prior to mailing. 37 CFR 1.10(b).

WARNING: Certificate of mailing (first class) or facsimile transmission procedures of 37 CFR 1.8 cannot be used to obtain a date of mailing or transmission for this correspondence.

EG130134065

NOTE: If one of the following 3 items apply, then complete and attach ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF A PRIOR U.S. APPLICATION CLAIMED and a NOTIFICATION IN PARENT APPLICATION OF THE FILING OF THIS CONTINUATION APPLICATION.

☐ Divisional.

☐ Continuation.

☐ Continuation-in-Part (C-I-P).

2. Benefit of Prior U.S. Application(s) (35 U.S.C. 119(e), 120, or 121)

NOTE: If the new application being transmitted is a divisional, continuation or a continuation-in-part of a parent case, or where the parent case is an International Application which designated the U.S., or benefit of a prior provisional application is claimed, then check the following item and complete and attach ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.

WARNING: *If an application claims the benefit of the filing date of an earlier filed application under 35 U.S.C. 120, 121 or 365(c), the 20-year term of that application will be based upon the filing date of the earliest U.S. application that the application makes reference to under 35 U.S.C. 120, 121 or 365(c). (35 U.S.C. 154(a)(2) does not take into account, for the determination of the patent term, any application on which priority is claimed under 35 U.S.C. 119, 365(a) or 365(b).) For a c-i-p application, applicant should review whether any claim in the patent that will issue is supported by an earlier application and, if not, the applicant should consider canceling the reference to the earlier filed application. The term of a patent is not based on a claim-by-claim approach. See Notice of April 14, 1995, 60 Fed. Reg. 20,195, at 20,205.*

WARNING: *When the last day of pendency of a provisional application falls on a Saturday, Sunday, or Federal holiday within the District of Columbia, any nonprovisional application claiming benefit of the provisional **must** be filed prior to the Saturday, Sunday or Federal holiday within the District of Columbia. See 37 C.F.R. § 1.78(a)(3).*

☐ The new application being transmitted claims the benefit of prior U.S. application(s) and enclosed are ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.

3. Papers Enclosed That Are Required For Filing Date Under 37 CFR 1.53(b) (Regular) or 37 CFR 1.153 (Design) Application

54 Pages of specification (including sequence listing)

3 Pages of claims

1 Pages of Abstract

8 Sheets of drawing

☑ formal

☐ informal

WARNING: ***DO NOT*** *submit original drawings. A high quality copy of the drawings should be supplied when filing a patent application. The drawings that are submitted to the Office must be on strong, white, smooth, and non-shiny paper and meet the standards according to § 1.84. If corrections to the drawings are necessary, they should be made to the original drawing and a high-quality copy of the corrected original drawing then submitted to the Office. Only one copy is required or desired. Comments on proposed new 37 CFR 1.84. Notice of March 9, 1988 (1990 O.G. 57-62).*

NOTE: "Identifying indicia, if provided, should include the application number or the title of the invention, inventor's name, docket number (if any), and the name and telephone number of a person to call if the Office is unable to match the drawings to the proper application. This information should be placed on the back of each sheet of drawing a minimum distance of 1.5 cm. (5/8 inch) down from the top of the page." 37 C.F.R. 1.84(c).

(complete the following, if applicable)

☐ The enclosed drawing(s) are photograph(s), and there is also attached a "PETITION TO ACCEPT PHOTOGRAPH(S) AS DRAWING(S)". 37 C.F.R. 1.84(b).

4. **Additional papers enclosed**

☑ Preliminary Amendment

☐ Information Disclosure Statement (37 CFR 1.98)

☐ Form PTO-1449

☐ Citations

☐ Declaration of Biological Deposit

☒ Submission of "Sequence Listing," computer readable copy and/or amendment pertaining thereto for biotechnology invention containing nucleotide and/or amino acid sequence.

☐ Authorization of Attorney(s) to Accept and Follow Instructions from Representative

☐ Special Comments

☐ Other

5. **Declaration or oath**

☐ Enclosed

executed by *(check **all** applicable boxes)*

☐ inventors.

☐ legal representative of inventors. 37 CFR 1.42 or 1.43

☐ joint inventor or person showing a proprietary interest on behalf of inventor who refused to sign or cannot be reached.

☐ This is the petition required by 37 CFR 1.47 and the statement required by 37 CFR 1.47 is also attached. *See item 13 below for fee.*

☑ Not Enclosed.

WARNING: *Where the filing is a completion in the U.S. of an International Application but where a declaration is not available or where the completion of the U.S. application contains subject matter in addition to the International Application the application may be treated as a continuation or continuation-in-part, as the case may be, utilizing ADDED PAGE FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION CLAIMED.*

☑ Application is made by a person authorized under 37 CFR 1.41(c) on behalf of *all the above named inventors.* (The declaration or oath, along with the surcharge required by 37 CFR 1.16(e) can be filed subsequently).

NOTE: It is important that all the correct inventor(s) are named for filing under 37 CFR 1.41(c) and 1.53(b).

☐ Showing that the filing is authorized. *(Not required unless called into question. 37 CFR 1.41(d).)*

6. **Inventorship Statement**

WARNING: *If the named inventors are each not the inventors of all the claims an explanation, including the ownership of the various claims at the time the last claimed invention was made, should be submitted.*

The inventorship for all the claims in this application are:

☑ The same

or

☐ Not the same. An explanation, including the ownership of the various claims at the time the last claimed invention was made,

☐ is submitted.

☐ will be submitted.

7. Language

NOTE: An application including a signed oath or declaration may be filed in a language other than English. A verified English translation of the non-English language application and the processing fee of $130.00 required by 37 CFR 1.17(k) is required to be filed with the application or within such time as may be set by the Office. 37 CFR 1.52(d).

NOTE: A non-English oath or declaration in the form provided or approved by the PTO need not be translated. 37 CFR 1.69(b).

☑ English

☐ non-English

☐ the attached translation is a verified translation. 37 CFR 1.52(d).

8. Assignment

☑ An assignment of the invention to MOGEN INTERNATIONAL NV

☐ is attached. A separate ☐ "COVER SHEET FOR ASSIGNMENT (DOCUMENT) ACCOMPANYING NEW PATENT APPLICATION" or ☐ FORM PTO 1595 is also attached.

☑ will follow.

NOTE: "If an assignment is submitted with a new application, send two separate letters—one for the application and one for the assignment." Notice of May 4, 1990 (1114 O.G. 77-78).

WARNING: *A newly executed "CERTIFICATE UNDER 37 CFR 3.73(b)" must be filed when a continuation-in-part application is filed by an assignee. Notice of April 30, 1993. 1150 O.G. 62-64.*

9. Certified Copy

Certified copy of application

Country	Appln. No.	Filed
Paraguay	9/96	January 12, 1996

from which priority is claimed

☐ is attached.

☑ will follow.

NOTE: The foreign application forming the basis for the claim for priority must be referred to in the oath or declaration. 37 CFR 1.55(a) and 1.63.

NOTE: This item is for any foreign priority for which the application being filed directly relates. If any parent U.S. application or International Application from which this application claims benefit under 35 U.S.C. 120 is itself entitled to priority from a prior foreign application then complete item 18 on the ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.

10. Fee Calculation (37 CFR 1.16)

A. ☑ Regular Application

Claims as Filed

Number Filed				Number Extra		Rate	Basic Fee 37 CFR 1.16(a) $770.00
Total Claims (37 CFR 1.16(c))	24	- 20	=	4	x $	22.00	88.00
Independent Claims (37 CFR 1.16(b))	4	- 3	=	1	x $	80.00	80.00
Multiple dependent claim(s), if any (37 CFR 1.16(d))					+ $	260.00	

☐ Amendment cancelling extra claims enclosed.

☑ Amendment deleting multiple-dependencies enclosed.

☑ Fee for extra claims is not being paid at this time.

NOTE: *If the fees for extra claims are not paid on filing they must be paid or the claims cancelled by amendment, prior to the expiration of the time period set for response by the Patent and Trademark Office in any notice of fee deficiency. 37 CFR 1.16(d).*

Filing Fee Calculation $ 770.00

B. ☐ Design application
($320.00 — 37 CFR 1.16(f))

Filing Fee Calculation $

C. ☐ Plant application
($530.00 — 37 CFR 1.16(g))

Filing Fee Calculation $

11. Small Entity Statement(s)

☐ Verified Statement(s) that this is a filing by a small entity under 37 CFR 1.9 and 1.27 is(are) attached.

Filing Fee Calculation (50% of **A**, **B** or **C** above) $

NOTE: *Any excess of the full fee paid will be refunded if a verified statement and a refund request are filed within 2 months of the date of timely payment of a full fee. 37 CFR 1.28(a).*

12. Request for International-Type Search (37 CFR 1.104(d)) ***(Complete, if applicable)***

☐ Please prepare an international-type search report for this application at the time when national examination on the merits takes place.

13. Fee Payment Being Made At This Time

☐ Not Enclosed

☐ No filing fee is to be paid at this time. *(This and the surcharge required by 37 CFR 1.16(e) can be paid subsequently.)*

☑ Enclosed

☑ basic filing fee $ 770.00

☐ Recording assignment
($40.00; 37 CFR 1.21(h)) (See attached "COVER SHEET FOR ASSIGNMENT ACCOMPANYING NEW APPLICATION.")

☐ Petition fee for filing by other than all the inventors or person on behalf of the inventor where inventor refused to sign or cannot be reached.
($130.00; 37 CFR 1.47 and 1.17(h)) $

☐ For processing an application with a specification in a non-English language.
($130.00; 37 CFR 1.52(d) and 1.17(k)) $

☐ Processing and retention fee
($130.00; 37 CFR 1.53(d) and 1.21(l))

☐ Fee for international-type search report
($40.00; 37 CFR 1.21(e)). $

NOTE: *37 CFR 1.21(l) establishes a fee for processing and retaining any application which is abandoned for failing to complete the application pursuant to 37 CFR 1.53(d) and this, as well as the changes to 37 CFR 1.53 and 1.78, indicate that in order to obtain the benefit of a prior U.S. application, either the basic filing fee must be paid or the processing and retention fee of §1.21(l) must be paid within 1 year from notification under §53(d).*

Total fees enclosed $ 770.00

14. Method of Payment of Fees

☑ Check in the amount of $ 770.00

☐ Charge Account No. 12-0425 in the amount of $

A duplicate of this transmittal is attached.

NOTE: *Fees should be itemized in such a manner that it is clear for which purpose the fees are paid. 37 CFR 1.22(b).*

15. Authorization to Charge Additional Fees

WARNING: *If no fees are to be paid on filing, the following items should not be completed.*

WARNING: *Accurately count claims, especially multiple dependent claims, to avoid unexpected high charges, if extra claim charges are authorized.*

☑ The Commissioner is hereby authorized to charge the following additional fees by this paper and during the entire pendency of this application to Account No. 12-0425.

☑ 37 CFR 1.16(a), (f) or (g) (filing fees)

☐ 37 CFR 1.16(b), (c) and (d) (presentation of extra claims)

NOTE: *Because additional fees for excess or multiple dependent claims not paid on filing or on later presentation must only be paid or these claims cancelled by amendment prior to the expiration of the time period set for response by the PTO in any notice of fee deficiency (37 CFR 1.16(d)), it might be best not to authorize the PTO to charge additional claim fees, except possibly when dealing with amendments after final action.*

☑ 37 CFR 1.16(e) (surcharge for filing the basic filing fee and/or declaration on a date later than the filing date of the application)

☑ 37 CFR 1.17 (application processing fees)

WARNING: *While 37 CFR 1.17(a), (b), (c) and (d) deal with extensions of time under §1.136(a), this authorization should be made only with the knowledge that: "Submission of the appropriate extension fee under 37 C.F.R. 1.136(a) is to no avail unless a request or petition for extension is filed." (Emphasis added). Notice of November 5,1985 (1060 O.G. 27)*

☑ 37 CFR 1.18 (issue fee at or before mailing of Notice of Allowance, pursuant to 37 CFR 1.311(b))

NOTE: Where an authorization to charge the issue fee to a deposit account has been filed before the mailing of a Notice of Allowance, the issue fee will be automatically charged to the deposit account at the time of mailing the notice of allowance. 37 CFR 1.311(b).

NOTE: 37 CFR 1.28(b) requires "Notification of any change in loss of entitlement to small entity status must be filed in the application ... prior to paying, or at the time of paying, ... issue fee". From the wording of 37 CFR 1.28(b): (a) notification of change of status must be made even if the fee is paid as "other than a small entity" and (b) no notification is required if the change is to another small entity.

16. Instructions As To Overpayment

☑ credit Account No. 12-0425

☐ refund

____
Signature of Attorney

Reg. No.

Tel. No.

WILLIAM R. EVANS
c/o LADAS & PARRY
26 WEST 61st STREET
NEW YORK, N.Y. 10023
Reg. No. 25,858 (212) 708-1946

☐ **Incorporation by reference of added pages**

(Check the following item if the application in this transmittal claims the benefit of prior U.S. application(s) (including an international application entering the U.S. stage as a continuation, divisional or C-I-P application) and complete and attach the ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED)

☐ Plus Added Pages for New Application Transmittal Where Benefit of Prior U.S. Application(s) Claimed

Number of pages added ___

☐ Plus Added Pages for Papers Referred to in Item 4 Above

Number of pages added ___

☐ Plus "Assignment Cover Letter Accompanying New Application"

Number of pages added ___

☑ **Statement Where No Further Pages Added**

(If no further pages form a part of this Transmittal, then end this Transmittal with this page and check the following item:)

☑ This transmittal ends with this page.

08/779460

MAIL ROOM
90 JAN 7 1997
PAT. & TRADEMARK OFF.

#1/2
Pre a

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Oscar Johannes Maria GODDIJN, et al.

For: ENHANCED ACCUMMULATION OF TREHALOSE IN PLANTS

Attorney Docket No.: U 011098-6

Commissioner of Patents and Trademarks
Washington, D.C. 20231

Sir:

PRELIMINARY AMENDMENT

Please amend the above identified application as follows:

IN THE CLAIMS

Claim 4, line 1, delete "any of claim 1 to 3" and replace therefor -- to claim 1--

Claim 6, line 1, delete "any one of claims 1 to 5" and replace therefor -- claim 1--

Claim 7, line 1, delete "any one of claims 1 to 5" and replace therefor -- claim 1--

Claim 8, line 1, delete "any one of claims 1 to 5" and replace therefor -- claim 1--

---

CERTIFICATE UNDER 37 CFR 1.10

I hereby certify that this paper is being deposited with the United States Postal Service on this date JANUARY 7, 1997 in an envelope as "EXPRESS MAIL POST OFFICE TO ADDRESSEE" Mailing Label Number EG130134065 addressed to the: Commissioner of Patents and Trademarks, Washington, D.C. 20231

GERALDINE MARTI
(Type or print name of person mailing paper)

Geraldine Marti
(Signature of person mailing paper)

NOTE: Each paper or fee referred to as enclosed herein has the number of the "EXPRESS MAIL" mailing label place thereon prior to mailing 37 CFR 1.16(b).

EG130134065

Claim 9, line 1, delete "any one of claims 1 to 8" and replace therefor -- claim 1--

Claim 10, line 2, delete "any one of the claims 1 to 9" and replace therefor -- claim 1--

Claim 18, line 2, delete "or 17"

Claim 24, line 3, delete "any one of claims 1 to 9" and replace therefor -- claim 1--

Respectfully submitted,

William R. Evans
Ladas & Parry
26 West 61st Street
New York, New York 10023
Reg.No.25858 (212) 708-1945

ENHANCED ACCUMULATION OF TREHALOSE IN PLANTS

FIELD OF THE INVENTION

The invention relates to a method for the production of trehalose in plant cells, and plants. The invention is particularly related to a method for increasing the levels of trehalose accumulation in plants by inhibiting the degradation of trehalose by trehalase. The invention further comprises higher plants, preferably *Angiospermae*, and parts thereof, which as a result of such methods, contain relatively high levels of trehalose. The invention further relates to plant cells, plants or parts thereof according to the invention obtained after processing thereof.

STATE OF THE ART

Trehalose is a general name given to D-glucosyl D-glucosides which comprise disaccharides based on two α-,α,ß- and ß,ß-linked glucose molecules. Trehalose, and especially α-trehalose alpha-D-glucopyranosyl(1-1)alpha-D-glucopyranoside is a widespread naturally occurring disaccharide. However, trehalose is not generally found in plants, apart from a few exceptions, such as the plant species *Selaginella lepidophylla (Lycophyta)* and *Myrothamnus flabellifolia*. Apart from these species, trehalose is found in root nodules of the *Leguminosae* (*Spermatophytae, Angiospermae)*, wherein it is synthesized by bacteroids; the trehalose so produced is capable of diffusing into the root cells. Apart from these accidental occurrences, plant species belonging to the *Spermatophyta* apparently lack the ability to produce and/or accumulate trehalose.

In International patent application WO 95/01446, filed on June 30, 1994 in the name of MOGEN International NV, a method is described for providing plants not naturally capable of producing trehalose with the capacity to do so.

In spite of the absence of trehalose as a substrate in most higher plant species, the occurrence of trehalose-degrading activity has been reported for a considerable number of higher plant species, including those known to lack trehalose. The responsible activity could be attributed to a trehalase enzyme.

Reports suggest that trehalose, when fed to plant shoots grown *in vitro* is toxic or inhibitory to the growth of plant cells (Veluthambi K. *et al.*, 1981, Plant Physiol. 68, 1369-1374). Plant cells producing low trehalase levels were found to be generally more sensitive to the adverse effects of trehalose, than plants exhibiting a higher level of trehalase activity. Trehalose-analogs, such as trehalose-amines were used to inhibit trehalase activity in shoots, making it possible to study the effects of trehalose fed to plant cells. Plant shoots which produce relatively high amounts of trehalase were adversely affected by the addition of trehalase inhibitors. Inhibition of trehalase activity in homogenates of callus and suspension culture of various *Angiospermae* using Validamycin is disclosed by Kendall *et al.*, 1990, Phytochemistry 29, 2525-2582.

It is an object of the present invention to provide plants and plant parts capable of producing and accumulating trehalose.

SUMMARY OF THE INVENTION

The invention provides a process for producing trehalose in plant cells capable of producing trehalase by growing plant cells having the genetic information required for the production of trehalose and trehalase, or cultivating a plant or a part thereof comprising such plant cells, characterised in that said plant cells are grown, or said plant or a part thereof, is cultivated in the presence of a trehalase inhibitor. Preferred plants or plant parts or plant cells have been genetically altered so as to contain a chimeric trehalose phosphate synthase gene in a plant expressible form. According to one embodiment said trehalose phosphate synthase gene comprises an open reading frame encoding trehalose phosphate synthase from *E. coli* in plant expressible form. More preferred is a gene coding for a bipartite enzyme with both trehalose phosphate synthase and trehalose phosphate phosphatase activities.

According to a further aspect of the invention, plants have been genetically altered so as to produce trehalose preferentially in certain tissues or parts, such as (micro-)tubers of potato. According to one embodiment the open reading frame encoding trehalose phosphate synthase from *E. coli* is downstream of the potato patatin promoter, to provide for

preferential expression of the gene in tubers and micro-tubers of *Solanum tuberosum*.

According to another aspect of the invention the plants are cultivated *in vitro*, for example in hydroculture.

According to another preferred embodiment said trehalase inhibitor comprises validamycin A in a form suitable for uptake by said plant cells, preferably in a concentration between 100 nM and 10 mM, preferably between 0.1 and 1 mM, in aqueous solution.

Equally suitable said trehalase inhibition can be formed by transformation of said plant with the antisense gene to a gene encoding the information for trehalase.

Also suitable as trehalase inhibitor is the 86 kD protein from the american cockroach (*Periplaneta americana*). This protein can be administered to a plant in a form suitable for uptake, and also it is possible that the plants are transformed with DNA coding for said protein.

The invention further provides plants and plant parts which accumulate trehalose in an amount above 0.01 % (fresh weight), preferably of a *Solanaceae* species, in particular *Solanum tuberosum* or *Nicotiana tabacum*, in particular a micro-tuber of *Solanum tuberosum* containing trehalose.

The invention also comprises the use of a plant, or plant part, according to the invention for extracting trehalose, as well as the use thereof in a process of forced extraction of water from said plant or plant part. According to yet another embodiment of the invention a chimaeric plant expressible gene is provided, comprising in sequence a transcription initiation region obtainable from a gene, preferentially expressed in a plant part, particularly the patatin gene from *Solanum tuberosum*, a 5'-untranslated leader, an open reading frame encoding a trehalose phosphate synthase activity, and downstream of said open reading frame a transcriptional terminator region.

According to yet another embodiment of the invention a chimaeric plant expressible gene is provided, comprising in sequence a transcription initiation region obtainable from a gene, preferentially expressed in a plant part, particularly the patatin gene from *Solanum tuberosum*, a 5'-untranslated leader, an open reading frame encoding a

trehalase coupled in the antisense orientation, and downstream of said open reading frame a transcriptional terminator region. A preferred plant expressible gene according to the invention is one wherein said transcriptional terminator region is obtainable from the proteinase inhibitor-II gene of *Solanum tuberosum*. The invention also provided vectors and recombinant plant genomes comprising a chimaeric plant expressible gene according to the invention, as well as a plant cell having a recombinant genome, a plant or a part thereof, consisting essentially of cells. A further preferred plant species according to this aspect is *Solanum tuberosum*, and a micro-tuber thereof.

The invention further provides a process for obtaining trehalose, comprising the steps of growing plant cells according to the invention or cultivating a plant according to the invention and extracting trehalose from said plant cells, plants or parts.

The following figures further illustrate the invention.

DESCRIPTION OF THE FIGURES

Figure 1. Schematic representation of binary vector pMOG845.

Figure 2. Schematic representation of multi-copy vector pMOG1192.

Figure 3. Alignments for maximal amino acid similarities of neutral trehalase from *S. cerevisiae* with periplasmatic trehalase from *E. coli*, small intestinal trehalase from rabbit and trehalase from pupal midgut of the silkworm, *Bombyx mori*. Identical residues among all trehalase enzymes are indicated in ***bold italics*** typeface. Conserved regions of the amino acid sequences were aligned to give the best fit. Gap's in the amino acid sequence are represented by dashes.
Positions of degenerated primers based on conserved amino acids are indicated by dashed arrows.

Figure 4. Alignment for maximal amino acid similarity of trehalases derived from *E. coli* (Ecoli2treh ; Ecolitreha), silkworm (Bommotreha), yellow mealworm (Tenmotreha), rabbit (Rabbitreha), *Solanum tuberosum* cv. Kardal (Potatotreha), and *S. cerevisiae* (Yeasttreha). Gap's in the amino

acid sequence are represented by dots.

Figure 5. Trehalase activity in leaf samples of *Nicotiana tabacum* cv. Samsun NN. Non-transgenic control plants are indicated by letters a-l, plants transgenic for pMOG1078 are indicated by numbers.

Figure 6. Trehalose accumulation in microtubers induced on stem segments derived from *Solanum tubersosum* cv. Kardal plants transgenic for both pMOG 845 (patatin driven $TPS_{E.coli}$ expression) and pMOG1027 (35SCaMV antisense-trehalase expression). N indicates the total number of transgenic lines screened. Experiments were performed in duplicate resulting in two values: a and b. ND: not determined.

DETAILED DESCRIPTION OF THE INVENTION

According to the present invention it has been found that the accumulation of an increased level of trehalose in plants and plant parts is feasible. This important finding can be exploited by adapting plant systems to produce and/or accumulate high levels of trehalose at lower cost.

According to one aspect of the invention the accumulation of increased levels of trehalose is achieved by inhibiting endogenous trehalases. Inhibition of trehalases can be performed basically in two ways: by administration of trehalase inhibitors exogenously, and by the production of trehalase inhibitors endogenously, for instance by transforming the plants with DNA sequences coding for trehalase inhibitors.

This inhibition can be equally well applied to plants which are transformed with enzymes which enable the production of trehalose, but also to plants which are able to synthesize trehalose naturally.

According to this first embodiment of the invention, trehalase inhibitors are administered to the plant system exogenously. Examples of trehalase inhibitors that may be used in such a process according to the invention are trehazolin produced in *Micromonospora*, strain SANK 62390 (Ando *et al.*, 1991, J. Antibiot. 44, 1165-1168), validoxylamine A, B, G, D-*gluco*-Dihydrovalidoxylamine A, L-*ido*-Dihydrovalidoxylamin A, Deoxynojirimycin (Kameda *et al.*, 1987, J. Antibiot. 40(4), 563-565), 5-

epi-trehazolin (Trehalostatin) (Kobayashi Y. *et al.*, 1994, J. Antiobiot. 47, 932-938), castanospermin (Salleh H.M. & Honek J.F. March 1990, FEBS 262(2), 359-362) and the 86kD protein from the american cockroach (*Periplaneta americana*) (Hayakawa *et al.*, 1989, J. Biol. Chem. 264(27), 16165-16169).

A preferred trehalase inhibitor according to the invention is validamycin A (1,5,6-trideoxy-3-o-ß-D-glucopyranosyl-5-(hydroxymethyl)-1-[[4,5,6-trihydroxy-3-(hydroxymenthyl)-2-cyclohexen-1-yl]amino]-D-chiro-inositol). Trehalase inhibitors are administered to plants or plant parts, or plant cell cultures, in a form suitable for uptake by the plants, plant parts or cultures. Typically the trehalase inhibitor is in the form of an aqueous solution of between 100 nM and 10 mM of active ingredient, preferably between 0.1 and 1 mM. Aqueous solutions may be applied to plants or plant parts by spraying on leaves, watering, adding it to the medium of a hydroculture, and the like. Another suitable formulation of validamycin is solacol, a commercially available agricultural formulation (Takeda Chem. Indust., Tokyo).

Alternatively, or in addition to using exogenously administered trehalase inhibitors, trehalase inhibitors may be provided by introducing the genetic information coding therefor. One form of such in-built trehalase inhibitor may consist of a genetic construct causing the production of RNA that is sufficiently complementary to endogenous RNA encoding for trehalase to interact with said endogenous transcript, thereby inhibiting the expression of said transcript. This so-called "antisense approach" is well known in the art (*vide* inter alia EP 0 240 208 A and the Examples to inhibit SPS disclosed in WO 95/01446).

A gene coding for trehalase has been isolated from a potato cDNA library and sequenced. The predicted amino acid sequence of trehalase as shown in SEQIDNO:10 is derived from the nucleotide sequence depicted in SEQIDNO: 9. A comparison of this sequence with known non-plant trehalase sequences learns that homology is scant. It is therefor questionable if such trehalase sequences used in an antisense approach are capable of inhibiting trehalase expression *in planta*.

Of course the most preferred embodiment of the invention is obtained by transforming a plant with the antisense trehalase gene which matches exactly with the endogenous trehalase gene. However, sequences

which have a high degree of homology can also be used. Thus, the antisense trehalase gene to be used for the transformation of potato will be directed against the nucleotide sequence depicted in SEQIDNO: 9. It is also demonstrated in this application that the potato trehalase sequence can also be used to inhibit trehalase expression in tomato since the potato sequence is highly homologous to the tomato trehalase sequence. Thus, it is envisaged that the potato sequence is usable at least in closely related species, but maybe also in other plants. This is even more the case, considering that it is usually enough to express only part of the homologous gene in the antisense orientation, in order to achieve effective inhibition of expression of the endogenous trehalase (*vide* Van der Krol *et al.*, 1990, Plant Molecular Biology, 14, 457-466). Furthermore, it is shown in this application that the potato trehalase sequence can be used for the detection of homology in other species.

Trehalase gene sequences of other plants can be elucidated using several different strategies. One of the strategies is to use the isolated potato cDNA clone as a probe to screen a cDNA library containing the cDNA of the desired plant species. Positive reacting clones can then be isolated and subcloned into suitable vectors.

A second strategy to identify such genes is by purifying the proteins which are involved in trehalose degradation. An example for such a strategy is the purification of a protein with acid invertase activity from potato (Solanum tuberosum L.) tubers (Burch *et al.*, Phytochemistry, Vol.31, No.6, pp. 1901-1904, 1992). The obtained protein preparation also exhibits trehalose hydrolysing activity. Disaccharide hydrolysing activity of protein preparations obtained after purification steps can be monitored as described by Dahlqvist (Analytical Biochemistry 7, 18-25, 1964).

After purifying the protein(s) with trehalose hydrolysing activity to homogeneity, the N-terminal amino acid sequence or the sequence of internal fragments after protein digestion is determined. These sequences enable the design of oligonucleotide probes which are used in a polymerase chain reaction (PCR) or hybridization experiments to isolate the corresponding mRNAs using standard molecular cloning techniques.

Alternatively, degenerated primers can be designed based on conserved sequences present in trehalase genes isolated from other

species. These primers are used in a PCR strategy to amplify putative trehalase genes. Based on sequence information or Southern blotting, trehalase PCR fragments can be identified and the corresponding cDNA's isolated.

An isolated cDNA encoding a trehalose degrading enzyme is subsequently fused to a promoter sequence in such a way that transcription results in the synthesis of antisense mRNA.

Another form of such an in-built trehalase inhibitor may consist of a genetic construct causing the production of a protein that is able to inhibit trehalase activity in plants. A proteinaceous inhibitor of trehalase has been isolated and purified from the serum of resting adult american cockroaches (*Periplaneta americana*) (Hayakawa et al., *supra*). This protein, of which the sequence partly has been described in said publication, can be made expressable by isolation of the gene coding for the protein, fusion of the gene to a suitable promoter, and transformation of said fused gene into the plant according to standard molecular biological methods.

A promoter may be selected from any gene capable of driving transcription in plant cells.

If trehalose accumulation is only desired in certain plant parts, such as potato (mini-)tubers, the trehalase inhibitory DNA construct (*e.g.* the antisense construct) comprises a promoter fragment that is preferentially expressed in (mini-)tubers, allowing endogenous trehalase levels in the remainder of the plant's cells to be substantially unaffected. Thus, any negative effects of trehalose to neighbouring plant cells due to trehalose diffusion, is counteracted by unaffected endogenous trehalase activity in the remainder of the plant.

In the Example illustrating the invention, wherein trehalose phosphate synthase is produced under the control of the patatin promoter fragment, also the trehalase-inhibitory construct may comprise a promoter fragment of the patatin gene.

Mutatis mutandis if trehalose is to be accumulated in tomato fruit, both a plant expressible trehalose phosphate synthase gene, which is at least expressed in the tomato fruit is to be used, as well as a plant expressible trehalase-inhibitory DNA construct, which should be expressed preferentially in the fruit, and preferably not, or not substantially,

outside the fruit. An example of a promoter fragment that may be used to drive expression of DNA-constructs preferentially in tomato fruit is disclosed in EP 0 409 629 A1. Numerous modifications of this aspect of the invention, that do not depart from the scope of this invention, are readily envisaged by persons having ordinary skill in the art to which this invention pertains.

An alternative method to block the synthesis of undesired enzymatic activity such as caused by endogenous trehalase is the introduction into the genome of the plant host of an additional copy of said endogenous trehalase gene. It is often observed that the presence of a transgene copy of an endogenous gene silences the expression of both the endogenous gene and the transgene (EP 0 465 572 A1).

According to one embodiment of the invention accumulation of trehalose is brought about in plants wherein the capacity of producing trehalose has been introduced by introduction of a plant expressible gene construct encoding trehalose phosphate synthase (TPS), see for instance WO 95/06126.

Any trehalose phosphate synthase gene under the control of regulatory elements necessary for expression of DNA in plant cells, either specifically or constitutively, may be used, as long as it is capable of producing active trehalose phosphate synthase activity. Most preferred are the trehalose phosphate synthase genes which also harbour a coding sequence for trehalose phosphate phosphatase activity, the so called bipartite enzymes. Such a gene, formerly only known to exist in yeast (see *e.g.* WO 93/17093), can also been found in most plants. This application describes the elucidation of such a gene from the sunflower *Helianthus annuus*, while also evidence is given for the existence of a homologous gene in *Nicotiana tabacum*. It is believed that the use of a bipartite enzyme enhances the production of trehalose because it enables completion of the metabolic pathway from UDP-glucose and glucose-6-phosphate into trehalose at one and the same site. Hence, the two-step synthesis is simplified into a one-step reaction, thereby increasing reaction speed and, subsequently, trehalose yield.

As genes involved in trehalose synthesis, especially genes coding for bipartite enzymes, become available from other sources these can be used in a similar way to obtain a plant expressible trehalose

synthesizing gene according to the invention.

Sources for isolating trehalose synthesizing activities include microorganisms (*e.g.* bacteria, yeast, fungi), but these genes can also be found in plants and animals.

The invention also encompasses nucleic acid sequences which have been obtained by modifying the nucleic acid sequence encoding enzymes active in the synthesis of trehalose by mutating one or more codons so that it results in amino acid changes in the encoded protein, as long as mutation of the amino acid sequence does not entirely abolish trehalose synthesizing activity.

According to another embodiment of the invention, plants are genetically altered to produce and accumulate trehalose in specific parts of the plant, which were selected on the basis of considerations such as substrate availability for the enzyme, insensitivity of the plant part to any putative adverse effects of trehalose on plant cell functioning, and the like. A preferred site for trehalose synthesising enzyme expression are starch storage parts of plants. In particular potato tubers are considered to be suitable plant parts. A preferred promoter to achieve selective enzyme expression in microtubers and tubers of potato is obtainable from the region upstream of the open reading frame of the patatin gene of potato (*Solanum tuberosum*).

Plants provide with a gene coding for trehalose phosphate synthase only may be further modified by introducing additional genes that encode phosphatases that are capable of the conversion of trehalose phosphate into trehalose. At least in potato tubers or micro-tubers, potato leaves and tobacco leaves and roots, endogenous phosphatase activity appears to be present, so that the introduction of a trehalose phosphate phosphatase (TPP) gene is not an absolute requirement.

Preferred plant hosts among the *Spermatophyta* are the *Angiospermae*, notably the *Dicotyledoneae*, comprising *inter alia* the *Solanaceae* as a representative family, and the *Monocotyledoneae*, comprising *inter alia* the *Gramineae* as a representative family. Suitable host plants, as defined in the context of the present invention include plants (as well as parts and cells of said plants) and their progeny which have been genetically modified using recombinant DNA techniques to cause or enhance production of trehalose in the desired plant or plant organ; these plants

may be used directly (*e.g.* the plant species which produce edible parts) in processing or the trehalose may be extracted and/or purified from said host. Crops with edible parts according to the invention include those which have flowers such as cauliflower (*Brassica oleracea*), artichoke (*Cynara scolymus*), fruits such as apple (*Malus, e.g. domesticus*), banana (*Musa, e.g. acuminata*), berries (such as the currant, *Ribes, e.g. rubrum*), cherries (such as the sweet cherry, *Prunus, e.g. avium*), cucumber (*Cucumis, e.g. sativus*), grape (*Vitis, e.g. vinifera*), lemon (*Citrus limon*), melon (*Cucumis melo*), nuts (such as the walnut, *Juglans, e.g. regia*; peanut, *Arachis hypogeae*), orange (*Citrus, e.g. maxima*), peach (*Prunus, e.g. persica*), pear (*Pyra, e.g. communis*), pepper (*Solanum, e.g. capsicum*), plum (*Prunus, e.g. domestica*), strawberry (*Fragaria, e.g. moschata*), tomato (*Lycopersicon, e.g. esculentum*), leafs, such as alfalfa (*Medicago sativa*), cabbages (such as *Brassica oleracea*), endive (*Cichoreum, e.g. endivia*), leek (*Allium porrum*), lettuce (*Lactuca sativa*), spinach (*Spinaciaoleraceae*), tobacco (*Nicotiana tabacum*), roots, such as arrowroot (*Maranta arundinacea*), beet (*Beta vulgaris*), carrot (*Daucus carota*), cassava (*Manihot esculenta*), turnip (*Brassica rapa*), radish (*Raphanus sativus*), yam (*Dioscorea esculenta*), sweet potato (*Ipomoea batatas*) and seeds, such as bean (*Phaseolus vulgaris*), pea (*Pisum sativum*), soybean (*Glycin max*), wheat (*Triticum aestivum*), barley (*Hordeum vulgare*), corn (*Zea mays*), rice (*Oryza sativa*), tubers, such as kohlrabi (*Brassica oleraceae*), potato (*Solanum tuberosum*), and the like. The edible parts may be conserved by drying in the presence of enhanced trehalose levels produced therein due to the presence of a plant expressible trehalose phosphate synthase gene.

The method of introducing the plant expressible gene coding for a trehalose-synthesizing enzyme, or any other sense or antisense gene into a recipient plant cell is not crucial, as long as the gene is expressed in said plant cell. The use of *Agrobacterium tumefaciens or Agrobacterium rhizogenes* - mediated transformation is preferred, but other procedures are available for the introduction of DNA into plant cells. Examples are transformation of protoplasts using the calcium/polyethylene glycol method, electroporation, microinjection and DNA-coated particle bombardment (Potrykus, 1990, Bio/Technol. **8**, 535-542). Also combinations of *Agrobacterium* and coated particle bombardment may be used. Also

transformation protocols involving other living vectors than *Agrobacterium* may be used, such as viral vectors (*e.g.* from the Cauliflower Mosaic Virus (CaMV) and or combinations of *Agrobacterium* and viral vectors, a procedure referred to as agroinfection (Grimsley N. *et al.*, 8 January 1987, Nature 325, 177-179). After selection and/or screening, the protoplasts, cells or plant parts that have been transformed are regenerated into whole plants, using methods known in the art (Horsch *et al.*, 1985, Science 225, 1229-1231).

The development of reproducible tissue culture systems for monocotyledonous crops, together with methods for introduction of genetic material into plant cells has facilitated transformation. Presently, preferred methods for transformation of monocot species are transformation with supervirulent *Agrobacterium*-strains, microprojectile bombardment of explants or suspension cells, and direct DNA uptake or electroporation (Shimamoto, *et al.*, 1989, Nature 338, 274-276). *Agrobacterium*-mediated transformation is functioning very well in rice (WO 94/00977). Transgenic maize plants have been obtained by introducing the *Streptomyces hygroscopicus bar*-gene, which encodes phosphinothricin acetyltransferase (an enzyme which inactivates the herbicide phosphinothricin), into embryogenic cells of a maize suspension culture by microprojectile bombardment (Gordon-Kamm, 1990, Plant Cell, 2, 603-618). The introduction of genetic material into aleurone protoplasts of other monocot crops such as wheat and barley has been reported (Lee, 1989, Plant Mol. Biol. 13, 21-30). Wheat plants have been regenerated from embryogenic suspension culture by selecting only the aged compact and nodular embryogenic callus tissues for the establishment of the embryogenic suspension cultures (Vasil, 1990 Bio/Technol. 8, 429-434).

Suitable DNA sequences for control of expression of the plant expressible genes (including marker genes), such as transcriptional initiation regions, enhancers, non-transcribed leaders and the like, may be derived from any gene that is expressed in a plant cell. Also intended are hybrid promoters combining functional portions of various promoters, or synthetic equivalents thereof. Apart from constitutive promoters, inducible promoters, or promoters otherwise regulated in their expression pattern, *e.g.* developmentally or cell-type specific, may be used to control expression of the plant expressible genes according to the

invention as long as they are expressed in plant parts that contain substrate for TPS.

To select or screen for transformed cells, it is preferred to include a marker gene linked to the plant expressible gene according to the invention to be transferred to a plant cell. The choice of a suitable marker gene in plant transformation is well within the scope of the average skilled worker; some examples of routinely used marker genes are the neomycin phosphotransferase genes conferring resistance to kanamycin (EP-B 131 623), the glutathion-S-transferase gene from rat liver conferring resistance to glutathione derived herbicides (EP-A 256 223), glutamine synthetase conferring upon overexpression resistance to glutamine synthetase inhibitors such as phosphinothricin (WO87/05327), the acetyl transferase gene from *Streptomyces viridochromogenes* conferring resistance to the selective agent phosphinothricin (EP-A 275 957), the gene encoding a 5-enolshikimate-3- phosphate synthase (EPSPS) conferring tolerance to N-phosphonomethylglycine, the *bar* gene conferring resistance against Bialaphos (*e.g.* WO 91/02071) and the like. The actual choice of the marker is not crucial as long as it is functional (*i.e.* selective) in combination with the plant cells of choice.

The marker gene and the gene of interest do not have to be linked, since co-transformation of unlinked genes (U.S. Patent 4,399,216) is also an efficient process in plant transformation.

Preferred plant material for transformation, especially for dicotyledonous crops are leaf-discs which can be readily transformed and have good regenerative capability (Horsch R.B. *et al.*, (1985) Science 227, 1229-1231).

It is immaterial to the invention how the presence of two or more genes in the same plant is effected. This can *inter alia* done be achieved by one of the following methods:

(a) transformation of the plant line with a multigene construct containing more than one gene to be introduced,

(b) co-transforming different constructs to the same plant line simultaneously,

(c) subsequent rounds of transformation of the same plant with the genes to be introduced,

(d) crossing two plants each of which contains a different gene to be

introduced into the same plant, or

(e) combinations thereof.

The field of application of the invention lies both in agriculture and horticulture, for instance due to improved properties of the modified plants as such (*e.g.* stress tolerance, such as cold tolerance, and preferably drought resistance, and increase in post-harvest quality and shelf-life of plants and plant products), as well as in any form of industry where trehalose is or will be applied in a process of forced water extraction, such as drying or freeze drying. Trehalose can be used or sold as such, for instance in purified form or in admixtures, or in the form of a plant product, such as a tuber, a fruit, a flower containing the trehalose, either in native state or in (partially) dehydrated form, and the like. Plant parts harbouring (increased levels of) trehalose phosphate or trehalose may be used or sold as such or processed without the need to add trehalose.

Also trehalose can be extracted and/or purified from the plants or plant parts producing it and subsequently used in an industrial process. In the food industries trehalose can be employed by adding trehalose to foods before drying. Drying of foods is an important method of preservation. Trehalose seems especially useful to conserve food products through conventional air-drying, and to allow for fast reconstitution upon addition of water of a high quality product (Roser *et al.*, July 1991, Trends in Food Science and Technology, pp. 166-169). The benefits include retention of natural flavors/fragrances, taste of fresh product, and nutritional value (proteins and vitamins). It has been shown that trehalose has the ability to stabilize proteins *e.g.* vaccines, enzymes and membranes, and to form a chemically inert, stable glass. The low water activity of such thoroughly dried food products prevents chemical reactions, that could cause spoilage.

Field crops like corn, cassava, potato, sugar beet and sugarcane have since long been used as a natural source for bulk carbohydrate production (starches and sucrose). The production of trehalose in such crops, facilitated by genetic engineering of the trehalose-biosynthetic pathway into these plant species, would allow the exploitation of such engineered crops for trehalose production.

Trehalose is also used in drying or storage of biological macromolecules, such as peptides, enzymes, polynucleotides and the like.

All references cited in this specification are indicative of the level of skill in the art to which the invention pertains. All publications, whether patents or otherwise, referred to previously or later in this specification are herein incorporated by reference as if each of them was individually incorporated by reference. In particular WO 95/01446, cited herein, describing the production of trehalose in higher plants by genetic manipulation is herein incorporated by reference.

The Examples given below illustrate the invention and are in no way intended to indicate the limits of the scope of the invention.

Experimental

DNA manipulations

All DNA procedures (DNA isolation from *E.coli*, restriction, ligation, transformation, etc.) are performed according to standard protocols (Sambrook *et al.* (1989) Molecular Cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, CSH, New York).

Strains

In all examples E.coli K-12 strain DH5α is used for cloning. The *Agrobacterium tumefaciens* strains used for plant transformation experiments are EHA 105 and MOG 101 (Hood *et al.* 1993, Trans. Research 2, 208-218)

Isolation of a patatin promoter/construction of pMOG546

A patatin promoter fragment is isolated from chromosomal DNA of *Solanum tuberosum* cv. Bintje using the polymerase chain reaction. A set of oligonucleotides, complementary to the sequence of the upstream region of the λpat21 patatin gene (Bevan, M., Barker, R., Goldsbrough, A., Jarvis, M., Kavanagh, T. and Iturriaga, G. (1986) Nucleic Acids Res. 14: 5564-5566), is synthesized consisting of the following sequences:

5′ AAG CTT ATG TTG CCA TAT AGA GTA G 3′	PatB33.2	(SEQIDNO:3)
5′ GTA GTT GCC ATG GTG CAA ATG TTC 3′	PatATG.2	(SEQIDNO:4)

These primers are used to PCR amplify a DNA fragment of 1123bp, using chromosomal DNA isolated from potato cv. Bintje as a template. The amplified fragment shows a high degree of similarity to the λpat21 patatin sequence and is cloned using EcoRI linkers into a pUC18 vector resulting in plasmid pMOG546.

Construction of pMOG 799

pMOG 799 harbours the TPS gene from *E. coli* under control of the double enhanced 35S Cauliflower Mosaic promoter. The construction of this binary vector is described in detail in International patent application WO 95/01446, incorporated herein by reference.

Construction of pMOG845.

Plasmid pMOG546 containing the patatin promoter is digested with NcoI-KpnI, incubated with *E. coli* DNA polymerase I in the presence of dATP and dCTP thereby destroying the NcoI and KpnI site and subsequently relegated. From the resulting vector a 1.1kb EcoRI-SmaI fragment containing the patatin promoter is isolated and cloned into pMOG798 (described in detail in WO 95/01446) linearized with SmaI-EcoRI consequently exchanging the 35S CaMV promoter for the patatin promoter. The resulting vector is linearized with HindIII and ligated with the following oligonucleotide duplex:

```
           (HindIII)  PstI     KpnI    HindIII

    5'    AGCT CTGCAG TGA GGTACC A      3'    TCV 11     (SEQIDNO:5)
    3'         GACGTC ACT CCATGG TTCGA  5'    TCV 12     (SEQIDNO:6)
```

After checking the orientation of the introduced oligonucleotide duplex, the resulting vector is linearized with PstI-HindIII followed by the insertion of a 950bp PstI-HindIII fragment harbouring the potato proteinase inhibitor II terminator (PotPiII) (An, G., Mitra, A., Choi, H.K., Costa, M.A., An, K., Thornburg, R. W. and Ryan, C.A. (1989) The Plant Cell 1: 115-122). The PotPiII terminator is isolated by PCR amplification using chromosomal DNA isolated from potato cv. Desiree as a template and the following set of oligonucleotides:

5'	GTACCCTGCAGTGTGACCCTAGAC	3'	TCV 15	(SEQIDNO:7)
5'	TCGATTCATAGAAGCTTAGAT	3'	TCV 16	(SEQIDNO:8)

The TPS expression cassette is subsequently cloned as a EcoRI-HindIII fragment into the binary vector pMOG402 resulting in pMOG845 (fig. 1). A sample of *E.coli* Dhα strain, harbouring pMOG845 has been deposited at the Centraal Bureau voor Schimmelcultures, Oosterstraat 1, P.O. Box 273, 3740 AG Baarn, The Netherlands, on January 4, 1995; the Accession Number given by the International Depositary Institution is CBS 101.95.

Triparental matings

The binary vectors are mobilized in triparental matings with the *E. coli* strain HB101 containing plasmid pRK2013 (Ditta G., Stanfield, S., Corbin, D., and Helinski, D.R. *et al.* (1980) Proc. Natl. Acad. Sci. USA 77, 7347) into *Agrobacterium tumefaciens* strain MOG101 or EHA105 and used for transformation.

Transformation of tobacco (*Nicotiana tabacum* SR1)

Tobacco is transformed by cocultivation of plant tissue with *Agrobacterium_tumefaciens* strain MOG101 containing the binary vector of interest as described. Transformation is carried out using cocultivation of tobacco (*Nicotiana tabacum* SR1) leaf disks as described by Horsch *et al.* 1985, Science 227, 1229-1231. Transgenic plants are regenerated from shoots that grow on selection medium containing kanamycin, rooted and transferred to soil.

Transformation of potato tuber discs

Potato (*Solanum tuberosum* cv. Kardal) is transformed with the *Agrobacterium* strain EHA 105 containing the binary vector of interest. The basic culture medium is MS30R3 medium consisting of MS salts (Murashige, T. and Skoog, F. (1962) Physiol. Plan. 14, 473), R3 vitamins (Ooms *et al.* (1987) Theor. Appl. Genet. 73, 744), 30 g/l sucrose, 0.5 g/l MES with final pH 5.8 (adjusted with KOH) solidified when necessary with 8 g/l Daichin agar. Tubers of *Solanum tuberosum* cv. Kardal are peeled and surface sterilized by burning them in 96% ethanol for 5 seconds.

Extinguish the flames in sterile water and cut slices of approximately 2 mm thickness. Disks are cut with a bore from the vascular tissue and incubated for 20 minutes in MS30R3 medium containing 1-5 x10^8 bacteria/ml of *Agrobacterium* EHA 105 containing the binary vector. Wash the tuber discs with MS30R3 medium and transfer them to solidified postculture medium (PM). PM consists of M30R3 medium supplemented with 3.5 mg/l zeatin riboside and 0.03 mg/l indole acetic acid (IAA). After two days, discs were transferred to fresh PM medium with 200 mg/l cefotaxim and 100 mg/l vancomycin. Three days later, the tuber discs are transferred to shoot induction medium (SIM) which consists of PM medium with 250 mg/l carbenicillin and 100 mg/l kanamycin. After 4-8 weeks, shoots emerging from the discs are excised and placed on rooting medium (MS30R3-medium with 100 mg/l cefotaxim, 50 mg/l vancomycin and 50 mg/l kanamycin). The shoots are propagated axenically by meristem cuttings.

<u>Potato stem-segment transformation protocol.</u>

Potato transformation experiments using stem-internodes were performed in a similar way as described by Newell C.A. *et al.*, Plant Cell Reports <u>10</u>: 30-34, 1990.

<u>Induction of micro-tubers</u>

Stem segments of *in vitro* potato plants harbouring an auxiliary meristem are transferred to micro-tuber inducing medium. Micro-tuber inducing medium contains 1 X MS-salts supplemented with R3 vitamins, 0.5 g/l MES (final pH= 5.8, adjusted with KOH) and solidified with 8 g/l Daishin agar, 60 g/l sucrose and 2.5 mg/l kinetin. After 3 to 5 weeks of growth in the dark at 24°C, micro-tubers are formed.

<u>Trehalose assay</u>

Trehalose was determined quantitatively by anion exchange chromatography with pulsed amperometric detection. Extracts were prepared by adding 1 ml boiling water to 1 g frozen material which was subsequently heated for 15′ at 100°C. Samples (25 µl) were analyzed on a Dionex DX-300 liquid chromatograph equipped with a 4 x 250 mm Dionex 35391 carbopac PA-1 column and a 4 x 50 mm Dionex 43096 carbopac PA-1 precolumn. Elution was with 100 mM NaOH at 1 ml/min. Sugars were detected with a pulsed

amperometric detector (Dionex, PAD-2). Commercially available trehalose (Sigma) was used as a standard.

Isolation of Validamycin A

Validamycin A is isolated from Solacol, a commercial agricultural formulation (Takeda Chem. Indust., Tokyo) as described by Kendall *et al.* (1990) Phytochemistry, Vol. 29, No. 8, pp. 2525-2528. The procedure involves ion exchange chromatography (QAE-Sephadex A-25 (Pharmacia), bed vol. 10 ml, equilibration buffer 0.2 mM Na-Pi pH 7) from a 3% agricultural formulation of Solacol. Loading 1 ml of Solacol on the column and eluting with water in 7 fractions, practically all Validamycin is recovered in fraction 4.

Based on a 100% recovery, using this procedure, the concentration of Validamycin A was adjusted to 110^{-3} M in MS-buffer, for use in trehalose accumulation tests.

Alternatively, Validamycin A and B may be purified directly from *Streptomyces hygroscopicus* var. *limoneus*, as described by Iwasa T. *et al.*, 1971, in The Journal of Antibiotics 24(2), 119-123, the content of which is incorporated herein by reference.

Construction of pMOG1027

pMOG1027 harbours the trehalase gene from *Solanum tuberosum* cv. Kardal in the reversed orientation under control of the double enhanced 35S Cauliflower Mosaic promoter. The construction of this vector is very similar to the construction of pMOG799 and can be performed by any person skilled in the art. After mobilization of this binary vector by triparental mating to *Agrobacterium*, this strain can be used to transform plant cells and to generate transgenic plants having reduced levels of trehalase activity.

Construction of pMOG1028

pMOG1028 harbours the trehalase gene from *Solanum tuberosum* cv. Kardal in the reversed orientation under control of the tuber specific patatin promoter. The construction of this vector is very similar to the construction of pMOG845 and can be performed by any person skilled in the

art. After mobilization of this binary vector by triparental mating to *Agrobacterium*, this strain can be used in potato transformation experiments to generate transgenic plants having reduced levels of trehalase activity in tuber-tissue.

Construction of pMOG 1078

To facilitate the construction of a binary expression cassette harbouring the trehalase cDNA clone in the "sense" orientation under control of the double enhanced 35S CaMV promoter, two HindIII sites were removed from the trehalase cDNA coding region (without changing the amino acid sequence) by PCR based point-mutations. In this way, a BamHI fragment was engineered that contained the complete trehalase open reading frame. This fragment was subsequently used for cloning in the binary vector pMOG800 behind the constitutive de35S CaMV promoter yielding pMOG1078. pMOG800 is derived from pMOG402; the KpnI site in the polylinker has been restored. pMOG402 is derived of pMOG23 (described in WO 95/01446) and harbours a restored neomycin phosphotransferase gene (Yenofsky R.L., Fine M., Pellow J.W., Proc Natl Acad Sci USA **87**: 3435-3439, 1990).

EXAMPLE 1

Trehalose production in tobacco plants transformed with pMOG799

Tobacco leaf discs are transformed with the binary vector pMOG799 using Agrobacterium tumefaciens. Transgenic shoots are selected on kanamycin. Transgenic plants are transferred to the greenhouse to flower and set seed after selfing (S1). Seeds of these transgenic plants are surface sterilised and germinated *in vitro* on medium with Kanamycin. Kanamycin resistant seedlings and wild-type tobacco plants are transferred to MS-medium supplemented with 10^{-3} M Validamycin A. As a control, transgenic seedlings and wild-type plants are transferred to medium without Validamycin A. Analysis of leaves and roots of plants grown on Validamycin A shows elevated levels of trehalose compared to the control plants (Table 1). No trehalose was detected in wild-type tobacco plants.

Table 1

	with Validamycin A		without Validamycin A	
	leaf	roots	leaf	roots
pMOG799.1	0.0081	0.0044	-	0.003
pMOG799.13	0.0110	0.0080	-	-
pMOG799.31	0.0008	0.0088	-	-
Wild-type SR1	-	-	-	-

EXAMPLE 2

Trehalose production in potato micro-tubers transformed with pMOG845

Potato *Solanum tuberosum* cv. Kardal tuber discs are transformed with *Agrobacterium tumefaciens* EHA105 harbouring the binary vector pMOG845. Transgenic shoots are selected on kanamycin. Micro-tubers (m-tubers) are induced on stem segments of transgenic and wild-type plants cultured on m-tuber inducing medium supplemented with 10^{-3} M Validamycin A. As a control, m-tubers are induced on medium without Validamycin A. M-tubers induced on medium with Validamycin A showed elevated levels of trehalose in comparison with m-tubers grown on medium without Validamycin A (Table 2). No trehalose was detected in wild-type m-tubers.

Table 2.

	Trehalose (% fresh weight)	
	+Validamycin A	-Validamycin A
845-2	0.016	-
845-4	-	-
845-8	0.051	-
845-13	0.005	-
845-22	0.121	-
845-25	0.002	-
wT Kardal	-	-

EXAMPLE 3

Trehalose production in hydrocultures of tobacco plants transformed with pMOG799

Seeds (S1) of selfed tobacco plants transformed with the binary vector pMOG799 are surface sterilised and germinated *in vitro* on MS20MS medium

containing 50 μg/ml Kanamycin. Kanamycin resistant seedlings are transferred to soil and grown in a growth chamber (temp. 23°C, 16 hours of light/day). After four weeks, seedlings were transferred to hydrocultures with ASEF clay beads with approximately 450 ml of medium. The medium contains 40 g/l Solacol dissolved in nano-water buffered with 0.5 g/l MES to adjust to pH 6.0 which is sieved through a filter to remove solid particles. Essential salts are supplemented by adding POKON™ (1.5 ml/l). The following antibiotics are added to prevent growth of micro-organisms: 500μg/ml Carbenicillin, 40μg/ml Nystatin and 100μg/ml Vancomycin. As a control, transgenic seedlings and wild-type plants are transferred to medium without Solacol. Analysis of leaves of plants grown on Solacol shows elevated levels of trehalose compared to the control plants (Table 3). No trehalose was detected in wild-type tobacco plants.

Table 3

	Solacol	Trehalose (%w/w)
pMOG 799.1-1	+	0.008
pMOG 799.1-2	+	0.004
pMOG 799.1-3	-	-
pMOG 799.1-4	-	-
pMOG 799.1-5	+	0.008
pMOG 799.1-6	-	-
pMOG 799.1-7	+	0.005
pMOG 799.1-8	-	-
pMOG 799.1-9	-	-
pMOG 799.1-10	+	0.007
Wild-type SR1-1	-	-
Wild-type SR1-2	+	-
Wild-type SR1-3	-	-
Wild-type SR1-4	+	-

Example 4

Cloning of a full length cDNA encoding trehalase from potato tuber

Using the amino acid sequence of the conserved regions of known trehalase genes (*E.coli*, Yeast, Rabbit, *B. mori*) (fig. 3), four degenerated primers were designed:

```
  C   C C  CGT    GT A   TTAT
GG GGI G TT   IGA  T TA    TGGGAC     Tase24 (SEQIDNO:11)
  T   A A  TAA    AG C   CGGC

                  TAA   GT
GTICCIGGIGGICGITT   IGA  T            Tase25 (SEQIDNO:12)
                  CGT   AG

     T   GA    TG    A  A
GGIGG TGI  ICGI  IAG TA TA            Tase26 (SEQIDNO:13)
     C   CT    CA    G  G

 C G  AT       A
I C TTI  CCATCC AAICCITC              Tase27 (SEQIDNO:14)
 G A  GC       G
```

Combinations of these primers in PCR experiments with genomic DNA and cDNA from *S. tuberosum* cv. Kardal leaf and tuber material respectively as template, resulted in several fragments of the expected length. A number of 190 bp. fragments obtained with the primer combination Tase24 and Tase 26 were subcloned into a pGEM T vector and sequenced. Several of the clones analyzed showed homology with known trehalase sequences. To exclude the isolation of non-plant derived trehalase sequences, Southern blot analysis was performed with gDNA from potato cv. Kardal. A number of clones isolated did not cross-hybridize with Kardal genomic DNA and were discarded. Two isolated clones were identical, gTase15.4 derived from a genomic PCR experiment and cTase5.2 derived from a PCR on cDNA, both showing hybridization in Southern blot analysis. One single hybridizing band was detected (EcoRI 1.5 Kb, HindIII 3 Kb and BamHI larger than 12 Kb) suggesting the presence of only one copy of the isolated PCR fragment.

A cDNA library was constructed out of poly A^+ RNA from potato tubers (cv. Kardal) using a Stratagene cDNA synthesis kit and the vector Lambda ZAPII. Recombinant phages (500.000) were screened with the radiolabeled cTase5.2 PCR fragment resulting in the identification of 3 positive clones. After purification, two clones were characterised with restriction enzymes revealing inserts of 2.15 and 2.3 kb respectively. Their nucleotide sequence was 100% identical. The nucleic acid sequence of one of these trehalase cDNA clones from *Solanum tuberosum* including

its open reading frame is depicted in SEQIDNO:9, while the aminoacid sequence derived from this nucleic acid sequence is shown in SEQIDNO:10. A plasmid harbouring an insert comprising the genetic information coding for trehalase has been deposited under no. CBS 804.95 with the Centraal Bureau voor Schimmelcultures, Oosterstraat 1, P.O. Box 273, 3740 AG Baarn, the Netherlands on December 8, 1995.

EXAMPLE 5

Homology between the trehalase gene from potato with other Solanaceae

Genomic DNA was isolated from tomato (Lycopersicon esculentum cv. Money maker), tobacco (Nicotiana tabacum cv. Petit havanna, SR1) and potato (Solanum tuberosum cv. Kardal), and subsequently digested with the restriction enzymes BamHI, BglII, NcoI, SpeI, AccI, HindIII and EcoRI. After gel-electrophoresis and Southern blotting, a [^{32}P]-alpha dCTP labelled trehalase potato cDNA probe was hybridized to the blot. Hybridization signals of almost similar strength were observed in the lanes with potato and tomato genomic DNA indicating a high degree of identity. Only a weak hybridization signal was observed in the lanes harbouring tobacco genomic DNA indicating a low degree of identity. A similar strategy can be used to identify trehalase genes from other crops and to select for crops were trehalase activity can be eliminated, via the anti-sense expression strategy, using a heterologous trehalase cDNA clone with sufficient homology. Alternatively, a homologous trehalase cDNA clone can be isolated and used in the anti-sense expression strategy.

EXAMPLE 6

Overexpression of a potato trehalase cDNA in *Nicotiana tabacum*

Tobacco leaf discs are transformed with the binary vector pMOG1078 using Agrobacterium tumefaciens. Transgenic shoots are selected on kanamycin and transferred to the greenhouse. Trehalase activity was determined in leaf samples of 26 transgenic and 12 non-transgenic control plants (Fig. 5). Trehalase activity up to ca. 17 μg trehalose/h/μg protein was measured compared to ca. 1 μg trehalose/h/μg protein for non-transgenic controls. This clearly confirms the identity of the potato trehalase cDNA.

EXAMPLE 7

Transformation of pMOG845 transgenic potato plants with pMOG1027

In order to super-transform pMOG845 transgenic potato lines with an anti-sense trehalase construct (pMOG1027), stem segments were cut from in vitro cultured potato shoots transgenic for pMOG845. Three parent lines were selected, pMOG845/11, /22 and /28 that revealed to accumulate trehalose in microtubers when grown on validamycin A. The stem segments were transformed with the binary vector pMOG1027 using Agrobacterium tumefaciens. Supertransformants were selected on Hygromycin and grown in vitro.

EXAMPLE 8

Trehalose production in tubers of potato plants transgenic for pMOG845 and pMOG1027

Microtubers were induced on explants of the pMOG845 transgenic potato plants supertransformed with pMOG1027 using medium without the trehalase inhibitor validamycin A. The accumulation of trehalose, up to 0.75 mg.g-1 fresh weight, was noted in the supertransformed lines proving the reduced trehalase activity in these lines using the anti-sense trehalase expression strategy (Fig. 6).

EXAMPLE 9

Isolation of a bipartite TPS/TPP gene from *Helianthus annuus*

To isolate a bipartite clone from H. annuus, a PCR amplification experiment was set up using two degenerate primers, TPS-deg2 and TPS-deg5. This primerset was used in combination with cDNA constructed on H. annuus leaf RNA as a template. A DNA fragment of approximately 650 bp. was amplified having a high similarity on amino acid level when compared to tps coding regions from E. coli and yeast. Based on its nucleotide sequence, homologous primers were designed and used in a Marathon RACE protocol (Clontech) to isolate the 5′ and 3′ parts of corresponding tps cDNA′s. Using primercombinations SUNGSP1(or 2)/AP1 in RACE PCR, no bands were observed whereas nested PCR with NSUNGSP1(or2)/AP2 resulted in several DNA fragments. Some of these fragments hybridized with a 32P labelled Sunflower tps fragment after Southern blotting. Two fragments of

circa 1.2 kb and 1.7 kb, corresponding respectively to the 5′ and 3′ part, were isolated from gel, subcloned and sequenced. The nucleotide sequence revealed a clear homology with known tps and tpp sequences indicating the bipartite nature of the isolated cDNA (SEQ ID NO 1). Using a unique XmaI site present in both fragments, a complete TPS/TPP bipartite coding region was obtained and subcloned in pGEM-T (Promega) yielding pMOG1192 (Fig. 2).

TPSdeg2:	tig git kit tyy tic aya yic cit tyc c	(SEQIDNO: 23)
TPSdeg5:	gyi aci arr ttc ati ccr tci c	(SEQIDNO: 27)

SUNGSP1:	cga aac ggg ccc atc aat ta	(SEQIDNO: 15)
SUNGSP2:	tcg atg aga tca atg ccg ag	(SEQIDNO: 16)
AP1 (Clontech):	cca tcc taa tac gac tca cta tag ggc	(SEQIDNO: 17)
NSUNGSP1:	cac aac agg ctg gta tcc cg	(SEQIDNO: 18)
NSUNGSP2:	caa taa cga act ggg aag cc	(SEQIDNO: 19)
AP2 (Clontech):	act cac tat agg gct cga gcg gc	(SEQIDNO: 20)

EXAMPLE 10

Isolation of a bipartite TPS/TPP gene from *Nicotiana tabacum*

Another strategy to isolate bipartite TPS/TPP genes from plants or other organisms involved the combined use of TPS and TPP primers in a single PCR reaction. As an example, a PCR was performed using cDNA generated on tobacco leaf total RNA and the primerset TPSdeg1 and TRE-TPP-16. Nested PCR, using the amplification mix of the first reaction as template, with TPSdeg2 and TRE-TPP-15 resulted in a DNA fragment of ca. 1.5 kb. Nested PCR of the original amplification mix with TPSdeg2 and TRE-TPP-10 yielded a DNA fragment of ca.1.2 kb.

Initial amplification using primer combination TPSdeg1 and TRE-TPP-6 followed by a nested PCR using primer combination TPSdeg2 and TRE-TPP-15 yielded a DNA fragment of ca. 1.5 kb.

Based on sequence analysis, the 1.2 kb and 1.5 kb amplified DNA fragments displayed a high degree of identity to TPS and TPP coding regions indicating that they encode a bipartite TPS/TPP proteins.

TPSdeg1:	GAY ITI ATI TGG RTI CAY GAY TAY CA	(SEQIDNO: 21)
TRE-TPP-16:	CCI ACI GTR CAI GCR AAI AC	(SEQIDNO: 22)
TPSdeg2:	TIG GIT KIT TYY TIC AYA YIC CIT TYC C	(SEQIDNO: 23)
TRE-TPP-15:	TGR TCI ARI ARY TCY TTI GC	(SEQIDNO: 24)
TRE-TPP-10:	CCR TGY TCI GCI SWI ARI CC	(SEQIDNO: 25)
TRE-TPP-6:	TCR TCI GTR AAR TCR TCI CC	(SEQIDNO: 26)

5+
28

#11

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS:

Oscar Johannes Maria GODDIJN
Teunis Cornelis VERWOERD
Ronny Wilhelmus Hermanus Henrika KRUTWAGEN
Eline VOOGD

(ii) TITLE OF INVENTION:

ENHANCED ACCUMULATION OF TREHALOSE IN PLANTS

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: LADAS & PARRY

(B) STREET: 26 WEST 61 STREET

(C) CITY: NEW YORK

(D) STATE: NY

(E) ZIP: 10023

(F) COUNTRY: USA

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3-1/4" Disk 1.44 MB

(B) COMPUTER: IBM PC Compatible

(C) OPERATING SYSTEM: Microsoft Windows for Workgroups 3.11

(D) SOFTWARE: WordPerfect 6.1 for Windows

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/779,460

(B) FILING DATE: 07-JAN-1997

(C) CLASSIFICATION: 435

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PY000009/96

(B) FILING DATE: 12-JAN-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: MASS, Clifford J.

(B) REGISTRATION NO.: 30,086

(C) REF./DOCKET NO.: U-011098-6

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE NUMBER: (212) 708-1890

(B) TELEAX NUMBER: (212)- 246-8959

(C) TELEX NUMBER: 233288

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2621 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 25..2485
(D) OTHER INFORMATION: /function= "trehalose phosph. synthase and trehalose phosph. phosphatase"
/product= "bipartite enzyme"

(ix) FEATURE:
(A) NAME/KEY: unsure
(B) LOCATION: 1609..1611

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
CTGATCCTGC GGTTTCATCA CAAT ATG ATA CTC TTA CAT CTG ATG CCC CTT        51
                           Met Ile Leu Leu His Leu Met Pro Leu
                            1               5

CAG ATG CTC CCA AAT AGG TTG ATT GTC GTA TCG AAT CAG TTA CCC ATA       99
Gln Met Leu Pro Asn Arg Leu Ile Val Val Ser Asn Gln Leu Pro Ile
 10                 15                  20                  25

ATC GCT AGG CTA AGA CTA ACG ACA ATG GAG GGT CCT TTT GGG ATT TCA      147
Ile Ala Arg Leu Arg Leu Thr Thr Met Glu Gly Pro Phe Gly Ile Ser
                30                  35                  40

CTT GGG ACG AGA GTT CGA TTT ACA TGC ACA TCA AAG ATG CAT TAC CCG      195
Leu Gly Thr Arg Val Arg Phe Thr Cys Thr Ser Lys Met His Tyr Pro
            45                  50                  55

CAG CCG TTG AGG TTT TCT ATT CTT GGC GAT CCA CTA AGG GCT GAC GTT      243
Gln Pro Leu Arg Phe Ser Ile Leu Gly Asp Pro Leu Arg Ala Asp Val
        60                  65                  70

GGC CCT ACC GAA CAA GAT GAC GTG TCA AAG ACA TTG CTC GAT AGG TTT      291
Gly Pro Thr Glu Gln Asp Asp Val Ser Lys Thr Leu Leu Asp Arg Phe
    75                  80                  85

AAT TGC GTT GCG GTT TTT GTC CCT ACT TCA AAA TGG GAC CAA TAT TAT      339
Asn Cys Val Ala Val Phe Val Pro Thr Ser Lys Trp Asp Gln Tyr Tyr
 90                 95                 100                 105

CAC TGC TTT TGT AAG CAG TAT TTG TGG CCG ATA TTT CAT TAC AAG GTT      387
His Cys Phe Cys Lys Gln Tyr Leu Trp Pro Ile Phe His Tyr Lys Val
                110                 115                 120

CCC GCT TCT GAC GTC AAG AGT GTC CCG AAT AGT CGG GAT TCA TGG AAC      435
Pro Ala Ser Asp Val Lys Ser Val Pro Asn Ser Arg Asp Ser Trp Asn
            125                 130                 135

GCT TAT GTT CAC GTG AAC AAA GAG TTT TCC CAG AAG GTG ATG GAG GCA      483
Ala Tyr Val His Val Asn Lys Glu Phe Ser Gln Lys Val Met Glu Ala
        140                 145                 150

GTA ACC AAT CGT AGC AAT TAT GTA TGG ATA CAT GAC TAC CAT TTA ATG      531
Val Thr Asn Arg Ser Asn Tyr Val Trp Ile His Asp Tyr His Leu Met
    155                 160                 165

ACG CTA CCG ACT TTC TTG AGG CGG GAT TTT TGT CGT TTT AAA ATC GGT      579
Thr Leu Pro Thr Phe Leu Arg Arg Asp Phe Cys Arg Phe Lys Ile Gly
170                 175                 180                 185

TTT TTT CTG CAT AGC CCG TTT CCT TCC TCG GAG GTT TAC AAG ACC CTA      627
Phe Phe Leu His Ser Pro Phe Pro Ser Ser Glu Val Tyr Lys Thr Leu
                190                 195                 200
```

```
CCA ATG AGA AAC GAG CTC TTG AAG GGT CTG TTA AAT GCT GAT CTT ATC        675
Pro Met Arg Asn Glu Leu Leu Lys Gly Leu Leu Asn Ala Asp Leu Ile
            205                 210                 215

GGG TTC CAT ACA TAC GAT TAT GCC CGT CAT TTT CTA ACG TGT TGT AGT        723
Gly Phe His Thr Tyr Asp Tyr Ala Arg His Phe Leu Thr Cys Cys Ser
        220                 225                 230

CGA ATG TTT GGT TTG GAT CAT CAG TTG AAA AGG GGG TAC ATT TTC TTG        771
Arg Met Phe Gly Leu Asp His Gln Leu Lys Arg Gly Tyr Ile Phe Leu
    235                 240                 245

GAA TAT AAT GGA AGG AGC ATT GAG ATC AAG ATA AAG GCG AGC GGG ATT        819
Glu Tyr Asn Gly Arg Ser Ile Glu Ile Lys Ile Lys Ala Ser Gly Ile
250                 255                 260                 265

CAT GTT GGT CGA ATG GAG TCG TAC TTG AGT CAG CCC GAT ACA AGA TTA        867
His Val Gly Arg Met Glu Ser Tyr Leu Ser Gln Pro Asp Thr Arg Leu
                270                 275                 280

CAA GTT CAA GAA GTC CAA AAA CGT TCG AAG GAA ATC GTG CTA CTG GGA        915
Gln Val Gln Glu Val Gln Lys Arg Ser Lys Glu Ile Val Leu Leu Gly
            285                 290                 295

GTT GAT GAT TTG GAT ATA TTC AAA GGT GTG AAC TTC AAG GTT TTA GCG        963
Val Asp Asp Leu Asp Ile Phe Lys Gly Val Asn Phe Lys Val Leu Ala
        300                 305                 310

TTG GAG AAG TTA CTT AAA TCA CAC CCG AGT TGG CAA GGG CGT GTG GAA       1011
Leu Glu Lys Leu Leu Lys Ser His Pro Ser Trp Gln Gly Arg Val Glu
    315                 320                 325

AAG GTG CAA ATC TTG AAT CCT CTG CGC CGT TGC CAA GAC GTC GAT GAG       1059
Lys Val Gln Ile Leu Asn Pro Leu Arg Arg Cys Gln Asp Val Asp Glu
330                 335                 340                 345

ATC AAT GCC GAG ATA AGA ACA GTC TGT GAA AGA ATC AAT AAC GAA CTG       1107
Ile Asn Ala Glu Ile Arg Thr Val Cys Glu Arg Ile Asn Asn Glu Leu
                350                 355                 360

GGA AGC CCG GGA TAC CAG CCC GTT GTG TTA ATT GAT GGG CCC GTT TCG       1155
Gly Ser Pro Gly Tyr Gln Pro Val Val Leu Ile Asp Gly Pro Val Ser
            365                 370                 375

TTA AGT GAA AAA GCT GCT TAT TAT GCT ATC GCC GAT ATG GCA ATT GTT       1203
Leu Ser Glu Lys Ala Ala Tyr Tyr Ala Ile Ala Asp Met Ala Ile Val
        380                 385                 390

ACA CCG TTA CGT GAC GGA CTG AAT CTT ATC CCG TAC GAG TAC GTC GTT       1251
Thr Pro Leu Arg Asp Gly Leu Asn Leu Ile Pro Tyr Glu Tyr Val Val
    395                 400                 405
```

TCC CGA CAA AGT GTT AAT GAC CCA AAT CCC AAT ACT CCA AAA AAG AGC 1299
Ser Arg Gln Ser Val Asn Asp Pro Asn Pro Asn Thr Pro Lys Lys Ser
410 415 420 425

ATG CTA GTG GTC TCC GAG TTC ATC GGT GTT TCA CTA TCT TTA ACC GGG 1347
Met Leu Val Val Ser Glu Phe Ile Gly Val Ser Leu Ser Leu Thr Gly
430 435 440

GCC ATA CGG GTC AAC CCA TGG GAT GAG TTG GAG ACA GCA GAA GCA TTA 1395
Ala Ile Arg Val Asn Pro Trp Asp Glu Leu Glu Thr Ala Glu Ala Leu
445 450 455

TAC GAC GCA CTC ATG GCT CCT GAT GAC CAT AAA GAA ACC GCC CAC ATG 1443
Tyr Asp Ala Leu Met Ala Pro Asp Asp His Lys Glu Thr Ala His Met
460 465 470

AAA CAG TAT CAA TAC ATT ATC TCC CAT GAT GTA GCT AAC TGG GCT AGC 1491
Lys Gln Tyr Gln Tyr Ile Ile Ser His Asp Val Ala Asn Trp Ala Ser
475 480 485

TTC TTT CAA GAT TTA GAG CAA GCG TGC ATC GAT CAT TCT CGT AAA CGA 1539
Phe Phe Gln Asp Leu Glu Gln Ala Cys Ile Asp His Ser Arg Lys Arg
490 495 500 505

TGC ATG AAT TTA GGA TTT GGG TTA GAT ACT AGA GTC GTC TTT TTG ATG 1587
Cys Met Asn Leu Gly Phe Gly Leu Asp Thr Arg Val Val Phe Leu Met
510 515 520

AGA AGT TTA GCA AGT TGG ATA AAG ATG TCT TGG AAG AAT GCT TAT TCC 1635
Arg Ser Leu Ala Ser Trp Ile Lys Met Ser Trp Lys Asn Ala Tyr Ser
525 530 535

ATG GCT CAA AAT CGG GCC ATA CTT TTG GAC TAT GAC GGC ACT GTT ACT 1683
Met Ala Gln Asn Arg Ala Ile Leu Leu Asp Tyr Asp Gly Thr Val Thr
540 545 550

CCA TCT ATC AGT AAA TCT CCA ACT GAA GCT GTT ATC TCC ATG ATC AAC 1731
Pro Ser Ile Ser Lys Ser Pro Thr Glu Ala Val Ile Ser Met Ile Asn
555 560 565

AAA CTG TGC AAT GAT CCA AAG AAC ATG GTG TTC ATC GTT AGT GGA CGC 1779
Lys Leu Cys Asn Asp Pro Lys Asn Met Val Phe Ile Val Ser Gly Arg
570 575 580 585

AGT AGA GAG AAA ATC TTG GCA GTT GGT TCG GCG CGT GTG AGA ACC CGC 1827
Ser Arg Glu Lys Ile Leu Ala Val Gly Ser Ala Arg Val Arg Thr Arg
590 595 600

CAT TGC ACT GAG CAC GGA TAC TTT ATA AGG TGG GCG GGT GAT CAA GAA 1875
His Cys Thr Glu His Gly Tyr Phe Ile Arg Trp Ala Gly Asp Gln Glu
605 610 615

TGG GAA ACG TGC GCA CGT GAG AAT AAT GTC GGG TGG ATG GAT GGA AAT 1923
Trp Glu Thr Cys Ala Arg Glu Asn Asn Val Gly Trp Met Asp Gly Asn
620 625 630

CTG AGG CCG GTT ATG AAT CTT TAT ACA GAA ACT ACT GAC GGT TCG TAT 1971
Leu Arg Pro Val Met Asn Leu Tyr Thr Glu Thr Thr Asp Gly Ser Tyr
635 640 645

ATT GAA AAG AAA GAA ACT GCA ATG GTT TGG CAC TAT GAA GAT GCT GAT 2019
Ile Glu Lys Lys Glu Thr Ala Met Val Trp His Tyr Glu Asp Ala Asp
650 655 660 665

AAA GAT CTT GGG TTG GAG CAG GCT AAG GAA CTG TTG GAC CAT CTT GAA 2067
Lys Asp Leu Gly Leu Glu Gln Ala Lys Glu Leu Leu Asp His Leu Glu
670 675 680

AAC GTG CTC GCT AAT GAG CCC GTT GGA GTG AAT CGA ACA GGT CAA TAC 2115
Asn Val Leu Ala Asn Glu Pro Val Gly Val Asn Arg Thr Gly Gln Tyr
685 690 695

ATT GTA GAA GTT AAA CCA CAG TCC CCC ATT AAT TAC CTT CTT GTT ATG 2163
Ile Val Glu Val Lys Pro Gln Ser Pro Ile Asn Tyr Leu Leu Val Met
700 705 710

ACA TTC ATA GGC ACT GAT TGT AGA ATC TTT AAC TTA AAT TTC TTT AAA 2211
Thr Phe Ile Gly Thr Asp Cys Arg Ile Phe Asn Leu Asn Phe Phe Lys
715 720 725

TAT GAA TGC AAT TAT AGG GGG TCA CTA AAA GGT ATA GTT GCA GAG AAG 2259
Tyr Glu Cys Asn Tyr Arg Gly Ser Leu Lys Gly Ile Val Ala Glu Lys
730 735 740 745

ATT TTT GCG TTC ATG GCT AAA AAG GGA AAA CAG GCT GAT TTC GTG TTG 2307
Ile Phe Ala Phe Met Ala Lys Lys Gly Lys Gln Ala Asp Phe Val Leu
750 755 760

ACG TTG AAT GAT AGA AGT GAT GAA GAC ATG TTT GTG GCC ATT GGG GAT 2355
Thr Leu Asn Asp Arg Ser Asp Glu Asp Met Phe Val Ala Ile Gly Asp
765 770 775

GGA ATA AAA AAG GGT CGG ATA ACT AAC AAC AAT TCA GTG TTT ACA TGC 2403
Gly Ile Lys Lys Gly Arg Ile Thr Asn Asn Asn Ser Val Phe Thr Cys
780 785 790

GTA GTG GGA GAG AAA CCG AGT GCA GCT GAG TAC TTT TTA AAT GAT GTC 2451
Val Val Gly Glu Lys Pro Ser Ala Ala Glu Tyr Phe Leu Asn Asp Val
795 800 805

TCG AGA AGC TCC GGG TGT CTC AGC AAC CAA GGA T GATCCGGAAG 2495
Ser Arg Ser Ser Gly Cys Leu Ser Asn Gln Gly
810 815 820

CTTCTCGTGA TCTTTATGAG TTAAAAGTTT TCGACTTTTT CTTCATCAAG ATTCATGGGA 2555

AAGTTGTTCA ATATGAACTT GTGTTCTTGG TTCTGGATTT TAGGGAGTCT ATGGATATAA 2615

CATTTC 2621

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 820 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ile Leu Leu His Leu Met Pro Leu Gln Met Leu Pro Asn Arg Leu
1 5 10 15

Ile Val Val Ser Asn Gln Leu Pro Ile Ile Ala Arg Leu Arg Leu Thr
20 25 30

Thr Met Glu Gly Pro Phe Gly Ile Ser Leu Gly Thr Arg Val Arg Phe
35 40 45

Thr Cys Thr Ser Lys Met His Tyr Pro Gln Pro Leu Arg Phe Ser Ile
50 55 60

Leu Gly Asp Pro Leu Arg Ala Asp Val Gly Pro Thr Glu Gln Asp Asp
65 70 75 80

Val Ser Lys Thr Leu Leu Asp Arg Phe Asn Cys Val Ala Val Phe Val
85 90 95

Pro Thr Ser Lys Trp Asp Gln Tyr Tyr His Cys Phe Cys Lys Gln Tyr
100 105 110

Leu Trp Pro Ile Phe His Tyr Lys Val Pro Ala Ser Asp Val Lys Ser
115 120 125

Val Pro Asn Ser Arg Asp Ser Trp Asn Ala Tyr Val His Val Asn Lys
130 135 140

Glu Phe Ser Gln Lys Val Met Glu Ala Val Thr Asn Arg Ser Asn Tyr
145 150 155 160

Val Trp Ile His Asp Tyr His Leu Met Thr Leu Pro Thr Phe Leu Arg
165 170 175

Arg Asp Phe Cys Arg Phe Lys Ile Gly Phe Phe Leu His Ser Pro Phe
180 185 190

Pro Ser Ser Glu Val Tyr Lys Thr Leu Pro Met Arg Asn Glu Leu Leu
195 200 205

Lys Gly Leu Leu Asn Ala Asp Leu Ile Gly Phe His Thr Tyr Asp Tyr
210 215 220

Ala Arg His Phe Leu Thr Cys Cys Ser Arg Met Phe Gly Leu Asp His
225 230 235 240

Gln Leu Lys Arg Gly Tyr Ile Phe Leu Glu Tyr Asn Gly Arg Ser Ile
245 250 255

Glu Ile Lys Ile Lys Ala Ser Gly Ile His Val Gly Arg Met Glu Ser
260 265 270

Tyr Leu Ser Gln Pro Asp Thr Arg Leu Gln Val Gln Glu Val Gln Lys
275 280 285

Arg Ser Lys Glu Ile Val Leu Leu Gly Val Asp Asp Leu Asp Ile Phe
290 295 300

Lys Gly Val Asn Phe Lys Val Leu Ala Leu Glu Lys Leu Leu Lys Ser
305 310 315 320

His Pro Ser Trp Gln Gly Arg Val Glu Lys Val Gln Ile Leu Asn Pro
325 330 335

Leu Arg Arg Cys Gln Asp Val Asp Glu Ile Asn Ala Glu Ile Arg Thr
340 345 350

Val Cys Glu Arg Ile Asn Asn Glu Leu Gly Ser Pro Gly Tyr Gln Pro
355 360 365

Val Val Leu Ile Asp Gly Pro Val Ser Leu Ser Glu Lys Ala Ala Tyr
370 375 380

Tyr Ala Ile Ala Asp Met Ala Ile Val Thr Pro Leu Arg Asp Gly Leu
385 390 395 400

Asn Leu Ile Pro Tyr Glu Tyr Val Val Ser Arg Gln Ser Val Asn Asp
405 410 415

Pro Asn Pro Asn Thr Pro Lys Lys Ser Met Leu Val Val Ser Glu Phe
420 425 430

Ile Gly Val Ser Leu Ser Leu Thr Gly Ala Ile Arg Val Asn Pro Trp
435 440 445

Asp Glu Leu Glu Thr Ala Glu Ala Leu Tyr Asp Ala Leu Met Ala Pro
450 455 460

Asp Asp His Lys Glu Thr Ala His Met Lys Gln Tyr Gln Tyr Ile Ile
465 470 475 480

Ser His Asp Val Ala Asn Trp Ala Ser Phe Phe Gln Asp Leu Glu Gln
485 490 495

```
Ala Cys Ile Asp His Ser Arg Lys Arg Cys Met Asn Leu Gly Phe Gly
            500                 505                 510

Leu Asp Thr Arg Val Val Phe Leu Met Arg Ser Leu Ala Ser Trp Ile
        515                 520                 525

Lys Met Ser Trp Lys Asn Ala Tyr Ser Met Ala Gln Asn Arg Ala Ile
    530                 535                 540

Leu Leu Asp Tyr Asp Gly Thr Val Thr Pro Ser Ile Ser Lys Ser Pro
545                 550                 555                 560

Thr Glu Ala Val Ile Ser Met Ile Asn Lys Leu Cys Asn Asp Pro Lys
                565                 570                 575

Asn Met Val Phe Ile Val Ser Gly Arg Ser Arg Glu Lys Ile Leu Ala
            580                 585                 590

Val Gly Ser Ala Arg Val Arg Thr Arg His Cys Thr Glu His Gly Tyr
        595                 600                 605

Phe Ile Arg Trp Ala Gly Asp Gln Glu Trp Glu Thr Cys Ala Arg Glu
    610                 615                 620

Asn Asn Val Gly Trp Met Asp Gly Asn Leu Arg Pro Val Met Asn Leu
625                 630                 635                 640

Tyr Thr Glu Thr Thr Asp Gly Ser Tyr Ile Glu Lys Lys Glu Thr Ala
                645                 650                 655

Met Val Trp His Tyr Glu Asp Ala Asp Lys Asp Leu Gly Leu Glu Gln
            660                 665                 670

Ala Lys Glu Leu Leu Asp His Leu Glu Asn Val Leu Ala Asn Glu Pro
        675                 680                 685

Val Gly Val Asn Arg Thr Gly Gln Tyr Ile Val Glu Val Lys Pro Gln
    690                 695                 700

Ser Pro Ile Asn Tyr Leu Leu Val Met Thr Phe Ile Gly Thr Asp Cys
705                 710                 715                 720

Arg Ile Phe Asn Leu Asn Phe Phe Lys Tyr Glu Cys Asn Tyr Arg Gly
                725                 730                 735

Ser Leu Lys Gly Ile Val Ala Glu Lys Ile Phe Ala Phe Met Ala Lys
            740                 745                 750

Lys Gly Lys Gln Ala Asp Phe Val Leu Thr Leu Asn Asp Arg Ser Asp
        755                 760                 765

Glu Asp Met Phe Val Ala Ile Gly Asp Gly Ile Lys Lys Gly Arg Ile
    770                 775                 780
```

Thr Asn Asn Asn Ser Val Phe Thr Cys Val Val Gly Glu Lys Pro Ser
785 790 795 800

Ala Ala Glu Tyr Phe Leu Asn Asp Val Ser Arg Ser Ser Gly Cys Leu
805 810 815

Ser Asn Gln Gly
820

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGCTTATGT TGCCATATAG AGTAG 25

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iii) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GTAGTTGCCA TGGTGCAAAT GTTC 24

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGCTCTGCAG TGAGGTACCA 20

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GACGTCACTC CATGGTTCGA 20

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTACCCTGCA GTGTGACCCT AGAC 24

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TCGATTCATA GAAGCTTAGA T 21

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2207 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Solanum tuberosum
(B) STRAIN: Kardal

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 161..1906

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 842..850
(D) OTHER INFORMATION: /function= "putative glycosylationsite"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CTTTTCTGAG TAATAACATA GGCATTGATT TTTTTTCAAT TAATAACACC TGCAAACATT 60

CCCATTGCCG GCATTCTCTG TTCTTACAAA AAAAAACATT TTTTTGTTCA CATAAATTAG 120

```
TTATGGCATC AGTATTGAAC CCTTTAACTT GTTATACAAT ATG GGT AAA GCT ATA      175
                                            Met Gly Lys Ala Ile
                                             1               5

ATT TTT ATG ATT TTT ACT ATG TCT ATG AAT ATG ATT AAA GCT GAA ACT      223
Ile Phe Met Ile Phe Thr Met Ser Met Asn Met Ile Lys Ala Glu Thr
                10                  15                  20

TGC AAA TCC ATT GAT AAG GGT CCT GTA ATC CCA ACA ACC CCT TTA GTG      271
Cys Lys Ser Ile Asp Lys Gly Pro Val Ile Pro Thr Thr Pro Leu Val
            25                  30                  35

ATT TTT CTT GAA AAA GTT CAA GAA GCT GCT CTT CAA ACT TAT GGC CAT      319
Ile Phe Leu Glu Lys Val Gln Glu Ala Ala Leu Gln Thr Tyr Gly His
        40                  45                  50

AAA GGG TTT GAT GCT AAA CTG TTT GTT GAT ATG TCA CTG AGA GAG AGT      367
Lys Gly Phe Asp Ala Lys Leu Phe Val Asp Met Ser Leu Arg Glu Ser
     55                  60                  65

CTT TCA GAA ACA GTT GAA GCT TTT AAT AAG CTT CCA AGA GTT GTG AAT      415
Leu Ser Glu Thr Val Glu Ala Phe Asn Lys Leu Pro Arg Val Val Asn
 70                  75                  80                  85

GGT TCA ATA TCA AAA AGT GAT TTG GAT GGT TTT ATA GGT AGT TAC TTG      463
Gly Ser Ile Ser Lys Ser Asp Leu Asp Gly Phe Ile Gly Ser Tyr Leu
                90                  95                 100

AGT AGT CCT GAT AAG GAT TTG GTT TAT GTT GAG CCT ATG GAT TTT GTG      511
Ser Ser Pro Asp Lys Asp Leu Val Tyr Val Glu Pro Met Asp Phe Val
            105                 110                 115

GCT GAG CCT GAA GGC TTT TTG CCA AAG GTG AAG AAT TCT GAG GTG AGG      559
Ala Glu Pro Glu Gly Phe Leu Pro Lys Val Lys Asn Ser Glu Val Arg
        120                 125                 130

GCA TGG GCA TTG GAG GTG CAT TCA CTT TGG AAG AAT TTA AGT AGG AAA      607
Ala Trp Ala Leu Glu Val His Ser Leu Trp Lys Asn Leu Ser Arg Lys
    135                 140                 145

GTG GCT GAT CAT GTA TTG GAA AAA CCA GAG TTG TAT ACT TTG CTT CCA      655
Val Ala Asp His Val Leu Glu Lys Pro Glu Leu Tyr Thr Leu Leu Pro
150                 155                 160                 165

TTG AAA AAT CCA GTT ATT ATA CCG GGA TCG CGT TTT AAG GAG GTT TAT      703
Leu Lys Asn Pro Val Ile Ile Pro Gly Ser Arg Phe Lys Glu Val Tyr
                170                 175                 180

TAT TGG GAT TCT TAT TGG GTA ATA AGG GGT TTG TTA GCA AGC AAA ATG      751
Tyr Trp Asp Ser Tyr Trp Val Ile Arg Gly Leu Leu Ala Ser Lys Met
            185                 190                 195
```

```
TAT GAA ACT GCA AAA GGG ATT GTG ACT AAT CTG GTT TCT CTG ATA GAT        799
Tyr Glu Thr Ala Lys Gly Ile Val Thr Asn Leu Val Ser Leu Ile Asp
        200                 205                 210

CAA TTT GGT TAT GTT CTT AAC GGT GCA AGA GCA TAC TAC AGT AAC AGA        847
Gln Phe Gly Tyr Val Leu Asn Gly Ala Arg Ala Tyr Tyr Ser Asn Arg
    215                 220                 225

AGT CAG CCT CCT GTC CTG GCC ACG ATG ATT GTT GAC ATA TTC AAT CAG        895
Ser Gln Pro Pro Val Leu Ala Thr Met Ile Val Asp Ile Phe Asn Gln
230                 235                 240                 245

ACA GGT GAT TTA AAT TTG GTT AGA AGA TCC CTT CCT GCT TTG CTC AAG        943
Thr Gly Asp Leu Asn Leu Val Arg Arg Ser Leu Pro Ala Leu Leu Lys
                250                 255                 260

GAG AAT CAT TTT TGG AAT TCA GGA ATA CAT AAG GTG ACT ATT CAA GAT        991
Glu Asn His Phe Trp Asn Ser Gly Ile His Lys Val Thr Ile Gln Asp
            265                 270                 275

GCT CAG GGA TCA AAC CAC AGC TTG AGT CGG TAC TAT GCT ATG TGG AAT       1039
Ala Gln Gly Ser Asn His Ser Leu Ser Arg Tyr Tyr Ala Met Trp Asn
        280                 285                 290

AAG CCC CGT CCA GAA TCG TCA ACT ATA GAC AGT GAA ACA GCT TCC GTA       1087
Lys Pro Arg Pro Glu Ser Ser Thr Ile Asp Ser Glu Thr Ala Ser Val
    295                 300                 305

CTC CCA AAT ATA TGT GAA AAA AGA GAA TTA TAC CGT GAA CTG GCA TCA       1135
Leu Pro Asn Ile Cys Glu Lys Arg Glu Leu Tyr Arg Glu Leu Ala Ser
310                 315                 320                 325

GCT GCT GAA AGT GGA TGG GAT TTC AGT TCA AGA TGG ATG AGC AAC GGA       1183
Ala Ala Glu Ser Gly Trp Asp Phe Ser Ser Arg Trp Met Ser Asn Gly
                330                 335                 340

TCT GAT CTG ACA ACA ACT AGT ACA ACA TCA ATT CTA CCA GTT GAT TTG       1231
Ser Asp Leu Thr Thr Thr Ser Thr Thr Ser Ile Leu Pro Val Asp Leu
            345                 350                 355

AAT GCA TTC CTT CTG AAG ATG GAA CTT GAC ATT GCC TTT CTA GCA AAT       1279
Asn Ala Phe Leu Leu Lys Met Glu Leu Asp Ile Ala Phe Leu Ala Asn
        360                 365                 370

CTT GTT GGA GAA AGT AGC ACG GCT TCA CAT TTT ACA GAA GCT GCT CAA       1327
Leu Val Gly Glu Ser Ser Thr Ala Ser His Phe Thr Glu Ala Ala Gln
    375                 380                 385

AAT AGA CAG AAG GCT ATA AAC TGT ATC TTT TGG AAC GCA GAG ATG GGG       1375
Asn Arg Gln Lys Ala Ile Asn Cys Ile Phe Trp Asn Ala Glu Met Gly
390                 395                 400                 405
```

65
42

```
CAA TGG CTT GAT TAC TGG CTT ACC AAC AGC GAC ACA TCT GAG GAT ATT      1423
Gln Trp Leu Asp Tyr Trp Leu Thr Asn Ser Asp Thr Ser Glu Asp Ile
                410                 415                 420

TAT AAA TGG GAA GAT TTG CAC CAG AAC AAG AAG TCA TTT GCC TCT AAT      1471
Tyr Lys Trp Glu Asp Leu His Gln Asn Lys Lys Ser Phe Ala Ser Asn
            425                 430                 435

TTT GTT CCG CTG TGG ACT GAA ATT TCT TGT TCA GAT AAT AAT ATC ACA      1519
Phe Val Pro Leu Trp Thr Glu Ile Ser Cys Ser Asp Asn Asn Ile Thr
        440                 445                 450

ACT CAG AAA GTA GTT CAA AGT CTC ATG AGC TCG GGC TTG CTT CAG CCT      1567
Thr Gln Lys Val Val Gln Ser Leu Met Ser Ser Gly Leu Leu Gln Pro
    455                 460                 465

GCA GGG ATT GCA ATG ACC TTG TCT AAT ACT GGA CAG CAA TGG GAT TTT      1615
Ala Gly Ile Ala Met Thr Leu Ser Asn Thr Gly Gln Gln Trp Asp Phe
470                 475                 480                 485

CCG AAT GGT TGG CCC CCC CTT CAA CAC ATA ATC ATT GAA GGT CTC TTA      1663
Pro Asn Gly Trp Pro Pro Leu Gln His Ile Ile Ile Glu Gly Leu Leu
                490                 495                 500

AGG TCT GGA CTA GAA GAG GCA AGA ACC TTA GCA AAA GAC ATT GCT ATT      1711
Arg Ser Gly Leu Glu Glu Ala Arg Thr Leu Ala Lys Asp Ile Ala Ile
            505                 510                 515

CGC TGG TTA AGA ACT AAC TAT GTG ACT TAC AAG AAA ACC GGT GCT ATG      1759
Arg Trp Leu Arg Thr Asn Tyr Val Thr Tyr Lys Lys Thr Gly Ala Met
        520                 525                 530

TAT GAA AAA TAT GAT GTC ACA AAA TGT GGA GCA TAT GGA GGT GGT GGT      1807
Tyr Glu Lys Tyr Asp Val Thr Lys Cys Gly Ala Tyr Gly Gly Gly Gly
    535                 540                 545

GAA TAT ATG TCC CAA ACG GGT TTC GGA TGG TCA AAT GGC GTT GTA CTG      1855
Glu Tyr Met Ser Gln Thr Gly Phe Gly Trp Ser Asn Gly Val Val Leu
550                 555                 560                 565

GCA CTT CTA GAG GAA TTT GGA TGG CCT GAA GAT TTG AAG ATT GAT TGC      1903
Ala Leu Leu Glu Glu Phe Gly Trp Pro Glu Asp Leu Lys Ile Asp Cys
                570                 575                 580

TAATGAGCAA GTAGAAAAGC CAAATGAAAC ATCATTGAGT TTTATTTTCT TCTTTTGTTA    1963

AAATAAGCTG CAATGGTTTG CTGATAGTTT ATGTTTTGTA TTACTATTTC ATAAGGTTTT    2023

TGTACCATAT CAAGTGATAT TACCATGAAC TATGTCGTTC GGACTCTTCA AATCGGATTT    2083

TGCAAAAATA ATGCAGTTTT GGAGAATCCG ATAACATAGA CCATGTATGG ATCTAAATTG    2143

TAAACAGCTT ACTATATTAA GTAAAAGAAA GATGATTCCT CTGCTTTAAA AAAAAAAAAA    2203
```

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 581 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
Met Gly Lys Ala Ile Ile Phe Met Ile Phe Thr Met Ser Met Asn Met
 1               5                  10                  15

Ile Lys Ala Glu Thr Cys Lys Ser Ile Asp Lys Gly Pro Val Ile Pro
            20                  25                  30

Thr Thr Pro Leu Val Ile Phe Leu Glu Lys Val Gln Glu Ala Ala Leu
        35                  40                  45

Gln Thr Tyr Gly His Lys Gly Phe Asp Ala Lys Leu Phe Val Asp Met
    50                  55                  60

Ser Leu Arg Glu Ser Leu Ser Glu Thr Val Glu Ala Phe Asn Lys Leu
65                  70                  75                  80

Pro Arg Val Val Asn Gly Ser Ile Ser Lys Ser Asp Leu Asp Gly Phe
                85                  90                  95

Ile Gly Ser Tyr Leu Ser Ser Pro Asp Lys Asp Leu Val Tyr Val Glu
            100                 105                 110

Pro Met Asp Phe Val Ala Glu Pro Glu Gly Phe Leu Pro Lys Val Lys
        115                 120                 125

Asn Ser Glu Val Arg Ala Trp Ala Leu Glu Val His Ser Leu Trp Lys
    130                 135                 140

Asn Leu Ser Arg Lys Val Ala Asp His Val Leu Glu Lys Pro Glu Leu
145                 150                 155                 160

Tyr Thr Leu Leu Pro Leu Lys Asn Pro Val Ile Ile Pro Gly Ser Arg
                165                 170                 175

Phe Lys Glu Val Tyr Tyr Trp Asp Ser Tyr Trp Val Ile Arg Gly Leu
            180                 185                 190

Leu Ala Ser Lys Met Tyr Glu Thr Ala Lys Gly Ile Val Thr Asn Leu
        195                 200                 205
```

```
Val Ser Leu Ile Asp Gln Phe Gly Tyr Val Leu Asn Gly Ala Arg Ala
    210                 215                 220

Tyr Tyr Ser Asn Arg Ser Gln Pro Pro Val Leu Ala Thr Met Ile Val
225                 230                 235                 240

Asp Ile Phe Asn Gln Thr Gly Asp Leu Asn Leu Val Arg Arg Ser Leu
                245                 250                 255

Pro Ala Leu Leu Lys Glu Asn His Phe Trp Asn Ser Gly Ile His Lys
            260                 265                 270

Val Thr Ile Gln Asp Ala Gln Gly Ser Asn His Ser Leu Ser Arg Tyr
        275                 280                 285

Tyr Ala Met Trp Asn Lys Pro Arg Pro Glu Ser Ser Thr Ile Asp Ser
    290                 295                 300

Glu Thr Ala Ser Val Leu Pro Asn Ile Cys Glu Lys Arg Glu Leu Tyr
305                 310                 315                 320

Arg Glu Leu Ala Ser Ala Ala Glu Ser Gly Trp Asp Phe Ser Ser Arg
                325                 330                 335

Trp Met Ser Asn Gly Ser Asp Leu Thr Thr Thr Ser Thr Thr Ser Ile
            340                 345                 350

Leu Pro Val Asp Leu Asn Ala Phe Leu Leu Lys Met Glu Leu Asp Ile
        355                 360                 365

Ala Phe Leu Ala Asn Leu Val Gly Glu Ser Ser Thr Ala Ser His Phe
    370                 375                 380

Thr Glu Ala Ala Gln Asn Arg Gln Lys Ala Ile Asn Cys Ile Phe Trp
385                 390                 395                 400

Asn Ala Glu Met Gly Gln Trp Leu Asp Tyr Trp Leu Thr Asn Ser Asp
                405                 410                 415

Thr Ser Glu Asp Ile Tyr Lys Trp Glu Asp Leu His Gln Asn Lys Lys
            420                 425                 430

Ser Phe Ala Ser Asn Phe Val Pro Leu Trp Thr Glu Ile Ser Cys Ser
        435                 440                 445

Asp Asn Asn Ile Thr Thr Gln Lys Val Val Gln Ser Leu Met Ser Ser
    450                 455                 460

Gly Leu Leu Gln Pro Ala Gly Ile Ala Met Thr Leu Ser Asn Thr Gly
465                 470                 475                 480

Gln Gln Trp Asp Phe Pro Asn Gly Trp Pro Pro Leu Gln His Ile Ile
                485                 490                 495
```

Ile Glu Gly Leu Leu Arg Ser Gly Leu Glu Glu Ala Arg Thr Leu Ala
500 505 510

Lys Asp Ile Ala Ile Arg Trp Leu Arg Thr Asn Tyr Val Thr Tyr Lys
515 520 525

Lys Thr Gly Ala Met Tyr Glu Lys Tyr Asp Val Thr Lys Cys Gly Ala
530 535 540

Tyr Gly Gly Gly Gly Glu Tyr Met Ser Gln Thr Gly Phe Gly Trp Ser
545 550 555 560

Asn Gly Val Val Leu Ala Leu Leu Glu Glu Phe Gly Trp Pro Glu Asp
565 570 575

Leu Lys Ile Asp Cys
580

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGYGGNMGMT TYRWNGARKT MTAYKRYTGG GAC 33

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(ix) FEATURE:
  (A) NAME/KEY: modified_base
  (B) LOCATION: 3
  (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
  (A) NAME/KEY: modified_base
  (B) LOCATION: 6
  (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
  (A) NAME/KEY: modified_base
  (B) LOCATION: 9
  (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
  (A) NAME/KEY: modified_base
  (B) LOCATION: 12
  (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
  (A) NAME/KEY: modified_base
  (B) LOCATION: 15
  (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
  (A) NAME/KEY: modified_base
  (B) LOCATION: 21
  (D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GTNCCNGGNG GNCGNTTYRW NGARKT 26

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 26 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(ix) FEATURE:

(A) NAME/KEY: modified_base
(B) LOCATION: 3
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 9
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 12
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 18
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGNGGYTGNS WNCGNYRNAG RTARTA 26

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 1
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 7
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 19
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 22
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

NSCRTTNRYC CATCCRAANC CNTC 24

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CGAAACGGGC CCATCAATTA 20

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TCGATGAGAT CAATGCCGAG 20

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CCATCCTAAT ACGACTCACT ATAGGGC 27

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CACAACAGGC TGGTATCCCG 20

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CAATAACGAA CTGGGAAGCC 20

23
50

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ACTCACTATA GGGCTCGAGC GGC 23

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 4
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 9
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GAYNTNATNT GGRTNCAYGA YTAYCA 26

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 3
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 12
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 18
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CCNACNGTRC ANGCRAANAC 20

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

45
52

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 2
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 5
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 8
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 14
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 20
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 23
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

TNGGNTKNTT YYTNCAYAYN CCNTTYCC 28

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 9
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 18
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

TGRTCNARNA RYTCYTTNGC 20

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 9
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 12
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 18
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CCRTGYTCNG CNSWNARNCC 20

77
54

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 17
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

TCRTCNGTRA ARTCRTCNCC 20

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 3
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 6
(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 15

(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION: 21
(D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GYNACNARRT TCATNCCRTC NC 22

79
85 56

CLAIMS

1. A process for producing trehalose in plant cells capable of producing trehalase by growing plant cells having the genetic information required for the production of trehalose and trehalase~~,~~ or cultivating a plant or a part thereof comprising ~~such~~ plant cells, characterised in that said plant cells are grown, or said plant ~~or a~~ part thereof, is cultivated in the presence of a trehalase inhibitor.

2. A process according to claim 1, wherein said plant cells have been genetically altered so as to contain a gene coding for a bipartite trehalose synthesizing enzyme in a plant expressible form.

3. A process according to claim 1, wherein said plant cells have been genetically altered so as to contain a chimeric trehalose phosphate synthase gene in a plant expressible form, preferably wherein the trehalose phosphate synthase gene comprises an open reading frame encoding trehalose phosphate synthase from *E. coli* in plant expressible form, more preferably wherein the open reading frame encoding trehalose phosphate synthase from *E. coli* is downstream of the CaMV 35S RNA promoter or the potato patatin promoter.

a

4. A process according ~~any of~~ claim 1 ~~to 3~~, wherein a *Solanum tuberosum* plant is cultivated, preferably wherein said plant has micro-tubers.

5. A process according to claim 4, wherein said plant is cultivated *in vitro*.

a

6. A process according to ~~any one of~~ claim~~s~~ 1 ~~to 5~~, wherein said trehalase inhibitor comprises validamycin A in a form suitable for uptake by said plant cells, said plant, or a part thereof. ~~thereof, preferably wherein the concentration of validamycin A is between 100 nM and 10 mM, more preferably between 0.1 and 1 mM, in aqueous solution.~~

D
D
D

EG130134065

7. A process according to ~~any one of~~ claim~~s~~ 1 ~~to 5~~, wherein said trehalase inhibitor comprises ~~the~~ 86kD protein of the cockroach (*Periplaneta americana*) in a form suitable for uptake by said plant cells, said plant, or a part thereof.

8. A process according to ~~any one of~~ claim~~s~~ 1 ~~to 5~~, wherein said plant cells have been genetically altered to contain the genetic information for a trehalase inhibitor~~, preferably wherein the trehalase inhibitor is the antisense gene to the gene encoding the information for trehalase or wherein the trehalase inhibitor is the 86kD protein of the American cockroach (*Periplaneta americana*)~~.

9. A process according to ~~any one of~~ claim~~s~~ 1 ~~to 8~~, wherein a plant, or a part thereof, accumulates trehalose in an amount greater than ~~above~~ 0.01 % (fresh weight).

10. A plant, or a part thereof, or plant cells, obtainable by a process according to ~~any one of the~~ claim~~s~~ 1 ~~to 9~~, which contain trehalose in an amount above 0.01% (fresh weight), preferably wherein said plant, or a part thereof is a *Solanaceae* species, more preferably *Solanum tuberosum* or *Nicotiana tabacum*.

11. A plant part according to claim 10, which is a tuber or a micro-tuber.

12. Tuber or micro-tubers of *Solanum tuberosum* containing trehalose.

13. Use of a plant, or plant part, according to claim 10 for extracting trehalose.

14. Use of a plant, or plant part, according to claim 10 in a process of forced extraction of water from said plant or plant part.

15. A ~~plant~~ according to claim 10, which has an increased stress tolerance~~, preferably increased drought tolerance~~.

16. A chimaeric plant expressible gene comprising in sequence a transcription initiation region obtainable from a gene, preferentially expressed in a plant part, particularly the patatin gene from *Solanum tuberosum*, a 5'-untranslated leader, an open reading frame encoding a trehalose phosphate synthase activity, and downstream of said open reading frame a transcriptional terminator region, preferably wherein said transcriptional terminator region is obtainable from the proteinase inhibitor-II gene of *Solanum tuberosum*.

17. A plant derived and plant expressible gene encoding a bipartite trehalose synthesizing enzyme.

a

18. A vector comprising a chimaeric plant expressible gene according to claim 16 ~~or 17~~.

19. A recombinant plant genome comprising a chimaeric gene according to claim 18.

20. A plant cell having a recombinant genome according to claim 18.

D

21. A plant or a part thereof, consisting essentially of cells according to claim 20, preferably a plant from the species *Solanum tuberosum*.

22. A plant part according to claim 21, which is a tuber or a micro-tuber.

23. A process for obtaining trehalose, comprising the steps of growing plant cells according to claim 20, or cultivating a plant according to claim 21, or cultivating a plant part according to any one of claims 21 or 22, extracting trehalose from said plant cells, plants or parts.

G

24. A process for obtaining trehalose, comprising the steps of producing trehalose in plant cells, a plant or a part thereof, according to a process of ~~any one of cl~~aims 1 ~~to 9~~, and separating or extracting trehalose from said plant cells, plant or part thereof.

a

Add D2

ADD F1 & G3